ASCAPE

Artificial intelligence Supporting CAncer Patients across Europe

Project Title	Artificial intelligence Supporting CAncer Patients across Europe
Project Acronym	ASCAPE
Grant Agreement No	875351
Instrument	Research and Innovation action
Call / Topic	H2020-SC1-DTH-2019 / Big data and Artificial Intelligence for monitoring health status and quality of life after the cancer treatment
Start Date of Project	01/01/2020
Duration of Project	36 months

D1.2 – ASCAPE Data Determinants and Piloting Validations

Work Package	WP1 – ASCAPE requirements, data structures, use cases and architecture specification
Lead Authors (Org)	Antonis Valachis (ORB), Imma Grau (FCRB), Montserrat Muñoz (FCRB), Raimundo Lozano (FCRB), Clara Amat (FCRB), Miltiadis Kokkonidis (INTRA)
Contributing Author(s) (Org)	Dimitris Miltiadou (UBITECH), Miltiadis Kokkonidis (INTRA), Lucian Itu (SIE), Anamaria Vizitiu (SIE), Radu Toev (SIE), Tudor Suditu (SIE), Mirjana Ivanović (UNSPMF), Vladimir Kurbalija (UNSPMF), Miloš Savić (UNSPMF), Brankica Bratić (UNSPMF), Manuel Perez (ATOS), Antonis Valachis (ORB), Imma Grau (FCRB), Montserrat Muñoz (FCRB), Raimundo Lozano (FCRB), Clara Amat (FCRB), Santiago Iriso (FCRB) Joan Grau-Sociats (FundiSYS), Angeles Fuentes (FundiSYS), Teresa Puig (FundiSYS), Thanos Kosmidis (CC), Paris Kosmidis (CC), Lazaros Tzelves (NKUA), Ioannis Varkarakis (NKUA)
Due Date	30.08.2020
Actual Date of Submission	30.08.2020
Version	1.0- ready for submission

The work described in this document has been conducted within the project ASCAPE. This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation programme under the Grant Agreement no 875351. This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of such content.

Dissemination Level

X	PU: Public (*on-line platform)
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Versioning and contribution history

Version	Date	Author	Notes
0.1	04.06.2020	Antonis Valachis (ORB), Imma Grau (FCRB), Montserrat Muñoz (FCRB), Raimundo Lozano (FCRB), Miltiadis Kokkonidis (INTRA), Angeles Fuentes (FundiSYS)	Initial Table of Contents
0.2	07.07.2020	Antonis Valachis (ORB)	Initial contribution on Pilot Data and QoL Data Determinants
0.3	09.07.2020	Dimitris Miltiadou (UBITECH), Miltiadis Kokkonidis (INTRA), Lucian Itu (SIE), Anamaria Vizitiu (SIE), Radu Toev (SIE), Tudor Suditu (SIE), Mirjana Ivanović (UNSPMF), Vladimir Kurbalija (UNSPMF), Miloš Savić (UNSPMF), Brankica Bratić (UNSPMF), Manuel Perez (ATOS)	Initial contributions on Wearables Selection, Retrospective and Prospective Data Collection, AI-Centric Evaluation, Health-Economics-Centric Evaluation
0.4	14.07.2020	Antonis Valachis (ORB), Imma Grau (FCRB), Montserrat Muñoz (FCRB), Raimundo Lozano (FCRB), Clara Amat (FCRB) Thanos Kosmidis (CC), Paris Kosmidis (CC), Lazaros Tzelves (NKUA), Ioannis Varkarakis (NKUA), Angeles Fuentes(FundiSYS), J. Grau-Sociats (FundiSYS), Serge Autexier (DFKI)	Initial contributions on Pilot Aims, Pilot Design, and Pilot Timelines
0.5	07.08.2020	All contributing authors	Updates to contributions and structure resulting in the first draft to be internally reviewed
0.6	19.08.2020	All contributing authors	Corrections to address issues identified during the first

			internal review
0.7	21.08.2020	Serge Autexier (DFKI), Johannes Rust (DFKI), Miltiadis Kokkonidis (INTRA)	Additional corrections and clarifications
0.8	28.08.2020	All contributing authors	Minor corrections to address issues identified during the second internal review
0.9	31.08.2020	Antonis Valachis (ORB), Imma Grau (FCRB), Montserrat Muñoz (FCRB), Raimundo Lozano (FCRB)	Task leaders' final check and minor corrections
1.0	31.08.2020	Miltiadis Kokkonidis (INTRA)	Work Package leader final check and minor corrections

Table of Contents

List of tables

List of figures

List of acronyms

AI	Artificial Intelligence
AMI	Adjusted Mutual Information
API	Application Programming Interface
BMI	Body-Mass-Index
DASH	Disabilities of Hand, Shoulder and Arm
DAM	Decision Analytic Modelling
DL	Deep Learning
DP	Differential Privacy
EHR	Electronic Health Record
EU	European Union
FCRB	Fundació Clínic Recerca Biomèdica (Clinic Foundation for Biomedical Research)
FDL	Federated Deep Learning
FL	Federated Learning
FFQ	Food Frequency Questionnaire
FIN	Fisher Network
GAD	General Anxiety Disorder
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
GPS	Global Positioning System
GPU	Graphics Processing Unit
HADS	Hospital Anxiety and Depression Scale
HCB	Hospital Clinic de Barcelona, third party of FCRB
HE	Homomorphic Encryption
HTTPS	HyperText Transfer Protocol Secure
ICT	Information and Communication Technology
IIEF	International Index of Erectile Function
IPSS	International Prostate Symptom Score
IT	Information Technology
KPI	Key Performance Indicator
LISAT	Life Satisfaction (Questionnaire)
LUTS	Lower Urinary Tract Symptoms
MAFEIP	Monitoring and Assessment Framework for the European Innovation Partnership on Active and Healthy Ageing
MAE	Mean Absolute Error
MCID	Minimal Clinical Importance Difference
MSE	Mean Squared Error
MI	Mutual information
ML	Machine Learning
NKUA	National and Kapodistrian University of Athens
NMI	Normalized Mutual Information
NYHA	New York Heart Association

ONTO-CR	ONTOlogies-based Clinical Repository (Barcelona Pilot)
OS	Operating System
PCC	Pearson Correlation Coefficient
PHQ	Patient Health Questionnaire
PREM	Patient Reported Experience Measurement
PROM	Patient Reported Outcome Measurement
PSA	Prostate Specific Antigen testing
RAE	Relative Absolute Error
RMSE	Root Mean Squared Error
RSE	Relative Squared Error
QoL	Quality of Life
RF	Radio Frequencies
SD	Standard deviation
SME	Small or Medium Enterprise
TAM	Technology Acceptance Model
UI	User Interface
UX	User Experience
VMs	Virtual Machines
WP	Work Package

Executive Summary

The ASCAPE pilots will provide the testing ground for the evaluation of ASCAPE technologies in a variety of settings including hospital-based settings, a primary healthcare setting, and a setting supporting patients remotely in parallel with their existing healthcare providers. A comprehensive evaluation framework, covering AI-centric evaluation of the algorithms, doctor-centric, patient-centric and health economics-centric evaluation, is outlined.

Patients will interact with ASCAPE indirectly. They will provide information that will help ASCAPE provide doctors more accurate prognostic models and more appropriate intervention recommendations. They will do so using a variety of means, including a specialised mobile application, a wearable device, web sites, email etc. Different pilots will use different data collection modalities reflecting future exploitation scenarios and offering opportunities for comparisons. Pilots will also adopt different approaches with regards to how doctors' recommendations are communicated to the patients, some focusing on direct contact with the doctor during patient visit, while others planning to utilise the technological means not only to collect input but also to provide doctor-approved advice to patients.

Doctors play a central role in ASCAPE as the project aims not to replace them, but to equip them with ICT tools to help them support their patients. They will be presented with ASCAPE predictions and intervention recommendations in a modern user interface (UI), designed to be able to provide quick overviews and the ability to drill down to obtain more detail when necessary. That UI will be provided by either the standalone ASCAPE Dashboard web application or by a new ASCAPE-powered version of the healthcare provider's IT system with the core ASCAPE Dashboard UI components integrated into it.

A major challenge ASCAPE will face is the fact that while pilots will need sufficient relevant data points from breast and prostate cancer patients in order to train the relevant ASCAPE AI models, the data collected for the numerous studies on the Quality of Life of cancer patients contacted are not, as a rule, available to third parties. In medical scientific literature, there is strong resistance to sharing non-aggregated patient data, making it difficult to build Big Data Machine Learning models. ASCAPE addresses this issue by technologies that allow building a body of AI knowledge without the individual patients' data being shared. Such AI knowledge will differ from information gained through scientific publications like original research articles or even systematic reviews and meta-analyses, in that it goes beyond general trends, being readily available and directly applicable to specific patient cases. The pilots will help build up such knowledge in the context of two data models, one for breast cancer and one for prostate cancer, and showcase ASCAPE's approach towards addressing the challenges of heterogenous data from different healthcare providers and data collection irregularities (patients not using a wearable for periods of time, not filling in questionnaires regularly etc.). Given the

challenges ahead, the success of pilots will make a very compelling case for the ability of ASCAPE to achieve its goals at scale.

1 Introduction

ASCAPE (Artificial intelligence Supporting CAncer Patients across Europe) is a H2020 Collaborative Research project involving 15 partners from 7 countries, including clinics, SMEs (small and medium-sized enterprises), research centres and universities. It falls under European Commission's Horizon 2020 Work Programme for 2018-2020 on "Health, demographic change and wellbeing''.

ASCAPE aims to take advantage of the recent advances in Big Data and AI (Artificial Intelligence) to support cancer patients' quality of life and health status after treatment.

The project will focus the training and application of its AI models on two types of cancer: breast and prostate, in order to achieve, within the confines of a 3-year collaborative research project:

a) sufficient representation of both genders and different age groups

b) measurable positive impact on quality of life and health status of patients.

ASCAPE aims to provide assistance to doctors, not to replace them. It will deliver AI-based predictions and medical intervention suggestions for QoL issues in breast and prostate cancer patients, with the express goal of offering tangible benefits for after-treatment quality of life improvements. The ASCAPE AI Models, based on large quantities of data, will provide to doctors (not only oncologists, but primary care practitioners as well) an additional source of knowledge, complementing existing knowledge from clinical studies, guidelines and clinical experience.

Knowledge captured in the ASCAPE AI Models will be readily applicable to individual patients cases for (a) providing predictions on their health status and (b) suggesting to doctors Quality-of-Life improving health interventions. This will be done with the help of a user-friendly interface allowing the doctor to explore different options and to visualise their effect.

Seeing the big picture, ASCAPE aims to minimise the gap between large- and small-sized healthcare providers in terms of their ability to use information technology infrastructure and data analysis approaches to their patients' advantage. The ASCAPE vision is to create (a) a true global Big Data Health AI platform widely accessible, with a state-of-the-art AI infrastructure and (b) privacy-preserving Big Data AI models trained on the basis of data collected by several healthcare providers, irrespective of their size and resources. Furthermore, ASCAPE aims at making these tools available to large- and small-sized healthcare providers alike, to primary care centres supporting cancer patients after cancer treatment, individual

doctors, researchers, SMEs and charities and any other party that could utilise them for the benefit of cancer patients around Europe and the world .

The ASCAPE pilots, in turn, aim to demonstrate ASCAPE's ability to deliver on that promise. In addition, the pilots will help inform and focus technical effort towards addressing medically relevant problems and shape the ASCAPE framework in a manner that ensures its applicability to the multitude of requirements emerging from the heterogeneity of the ASCAPE pilots.

Section 2 describes the aims and the design of the pilots. Section 3 focuses on data determinants in terms of QoL issues to be predicted and manage accordingly as well as the variables of interest in pre-existing and in prospective datasets that will be used for providing predictions and intervention suggestions. Section 4 lays down the pilots' evaluation methodology, covering four different perspectives: AI-centric, doctor-centric, patient-centric, and health economics-centric evaluation of the pilots. Finally, Section 5 provides the core conclusions of the deliverable and links the work presented herein with the following tasks of WP1 and the other work packages.

2 Pilots Design

2.1 Aims of the Pilots

The four ASCAPE pilots will help ASCAPE develop and test AI capabilities addressing QoL issues of breast and prostate cancer patients.

2.1.1 Athens Pilot Aims

The Urology Department of Sismanoglio Hospital is part of the School of Medicine of the National and Kapodistrian University of Athens (NKUA). During the last 5 years more than 20 clinical studies took place in the Department, most of which are mainly observational, while a minority comprise controlled interventional trials. The Athens ASCAPE trial is the first of its kind for NKUA in terms of using AI technology for assessing and potentially improving patients' QoL.

Participation in ASCAPE is a first-class opportunity for the Urology Department of Athens Pilot to move from antiquated practices of paper-based patient record keeping to a cutting-edge AI-powered system that will not only assist doctors with find relevant information faster but also help keep track of patients' QoL issues despite staff shortages, in line with ASCAPE's promise of challenging the Iron Triangle of Health orthodoxy. This transformation will be put in practice and to the test in the context of the ASCAPE Athens prostate cancer study. Under certain conditions, the transformation will not be an ephemeral one ending with the project's completion, but one leading to long term benefits for the Urology Department of Sismanoglio Hospital and its patients. Furthermore, the data collected can be an asset for pooled studies in the future.

NKUA fully supports ASCAPE's aims of the formation of large networks of healthcare providers contributing to the collaborative building up of Big Data AI models, with the aim to democratising knowledge and raising the level of provided healthcare services for patients not only in Greece, but in other countries also.

The ASCAPE trial will also give the opportunity to discover ''hidden'' associations between several clinical parameters and their effect on QoL and also the effectiveness of several interventions in improving such issues. These associations cannot usually be detected using the conventional statistical methods.

2.1.2 Barcelona Pilot Aims

The collaboration in the Barcelona pilot is led by researchers from the Hospital Clínic de Barcelona (HCB) through the FCRB and the iSYS Foundation, which provides the

digital platform Xemio. The pilot will also include other entities such as Primary Care Centres in Barcelona.

The objective of the pilot is to improve the quality of life of Breast cancer survivors. As a result, the main aim of the Barcelona pilot is to enhance the services to cancer patients, by allowing patients to benefit from the learnings of the central ASCAPE system. By participating in this project, HCB with the collaboration of FundiSYS will continue to be a leader in ICT targeting patient experience.

Along with ASCAPE, the Barcelona pilot will also use two recent ICT innovations developed between these two institutions to improve patient outcomes and experience: the Xemio app, and ONTO-CR. The secondary aim of the Barcelona pilot is to introduce the app Xemio to breast cancer patients. Xemio is an ICT that will improve the self-management of the disease by the patients and will also improve management of the disease by the physicians because Xemio will be linked to ONTO-CR. The incorporation of Xemio into the health record system will be done with ONTO-CR, which is a semantically interoperable clinical repository, developed by HCB researchers, based on ontologies and conforming to the CEN/ISO 13606 standard. Likewise, ASCAPE will also be integrated with ONTO-CR, thus making ONTO-CR the focal point of data collection and doctors' access to patient data in the Barcelona pilot. ASCAPE, XEMIO and ONTO-CR will greatly advance the introduction of ICT's in the hospital and primary care system with the aim of improving the quality of life of breast cancer patients.

2.1.3 CareAcross Pilot Aims

CareAcross is a company that operates platforms offering personalised support for cancer patients. This support is provided directly to patients, who self-register, and engage with, these platforms on their own. As such, this pilot is significantly different from the other pilots, which are conducted by actual clinics (and primary care centres in the Barcelona pilot's case). Therefore, the CareAcross ASCAPE pilot is quite well aligned with the aims of democratising care and expanding remote monitoring and improvement of health status and quality of life beyond the confines of a clinical setting.

At the same time, it is understandable that such interactions are not always possible without a clinical examination. After all, CareAcross is not aiming to replace the hospital or clinicians. On the contrary, it aims to improve the relationship between the patients and their clinicians, without requiring their constant interaction and without disrupting the communication modalities that may be inherent in various healthcare systems.

As a result, the main aim of the CareAcross pilot is to expand the scope and enhance the level of the CareAcross services to cancer patients, by allowing patients to benefit from the learnings of the central ASCAPE system while not requiring them to be attached to a specific clinic or hospital. Experience gained from

the pilot and participation in the ASCAPE project will subsequently enable CareAcross to explore ways to expand its services further, by introducing more custom remote monitoring capabilities.

Within the context of ASCAPE, the CareAcross pilot will address both breast cancer and prostate cancer patients.

2.1.4 Örebro Pilot Aims

Örebro pilot will be led by the Department of Oncology, Örebro University Hospital which is an academic Department offering cutting-edge cancer care in patients with cancer. The Department has a dedicated Clinical Trial Unit in order to incorporate clinical research into daily clinical practice. The Örebro pilot is, therefore, an example on how ASCAPE can be developed and integrated in a clinical setting of a tertiary Hospital with existing electronic medical records and follow-up protocols for cancer patients. Specifically, the Department of Oncology is responsible for the follow-up of patients with breast cancer after curative treatment as well as patients with prostate cancer after radiotherapy as curative therapy and both groups will be included in the Örebro pilot. For the breast cancer cohort, the Department of Oncology, Uppsala University Hospital will also include patients to the pilot

The Örebro pilot has two aims: the primary aim is to evaluate the QoL aspects in patients with breast and prostate cancer treatment with curative intention during the 1st year of diagnosis and to facilitate, through structured data collection, the development of AI-based models predicting potential impairment in QoL aspects through the ASCAPE.

The secondary aims of the pilot are to implement the ASCAPE-based platform for predictions and interventions in a real-world clinical setting and investigate the clinicians' views and experiences in using AI-based models in clinical practice and patients' experiences in the concept of AI-based follow-up.

2.2 Breast cancer studies

Three of the four pilot will engage in studies focusing on breast cancer patients' QoL, the Barcelona pilot led by HCB, FCRB and Fundación iSYS, the Orebro pilot led by Orebro University (with the collaboration of Uppsala University) and the CareAcross pilot led by CareAcross.

2.2.1 Barcelona breast cancer study

2.2.1.1 Study design, population and geographic coverage

The Barcelona pilot is designed as a longitudinal study. Breast cancer survivors that have finished treatment, are in follow-up with HCB or their primary care doctor will be included in this trial. Given the necessity to first, feed data to the ASCAPE algorithms

as well as the need of a control group, some of the participants from the Barcelona pilot will complete the intervention as controls and will not receive the ASCAPE predictions, and the rest will participate in the intervention and will receive the benefits from the ASCAPE predictions. The intervention will be 12 months long for each participant. There will be a total of 150 participants.

In addition, the Barcelona pilot will also study the effect of the use of the Xemio app on quality of life. To that effect, participants will be randomised into two arms: a group without the Xemio app and an intervention group that will use the Xemio app. Half will use the Xemio app and half will not use the Xemio app.

Patients of a pre-existent cohort in Hospital Clinic will be invited to participate in the trial by their treating physician. Currently, there are ~200 patients in this cohort and all of them will be invited. Following the informed consent process and signature, each participant will be randomised into one of the two groups.

During the duration of the study the trial participants will answer QoL questionnaires during the initial visit, at 3 months, 6 months, 9 months, and 12 months. Participants with the Xemio app will be able to answer these questionnaires on the app itself. Participants on the control group will answer the QoL questionnaires on a secure web form on Onto-CR.

Inclusion criteria

1. Female patients >18 years old
2. Previous diagnosis of breast cancer, without active disease at the moment of inclusion and in follow-up (surgery, chemotherapy for early disease and radiotherapy must be completed). >1 year post-chemotherapy treatment or >1 year post-surgery in case that chemotherapy was not required.
3. Signature of the informed consent form.
4. Ownership of a smartphone and knowledge on how to use phone applications.
5. Belong to the covered geographical population.
6. Being treated by a collaborating physician within the geographical area.

Exclusion criteria

1. Acute cancer process at the moment of recruitment.
2. Having important comorbidities such as: heart failure NYHA class 4, unable to leave her residence, neurodegenerative disease, chronic obstructive pulmonary disease, ...
3. Inability to give informed consent.

The geographic area that will be covered by the Barcelona pilot has more than 500.000 inhabitants. It is composed of 20 Primary Health Care Centers, 2 smaller Hospital and HCBl as Reference Center. In addition, this pilot will benefit from having a diverse population. The southern part of this area has an average gross domestic product (GDP) of 24,000 - 30,0000€ and the northern part of the area has a GDP of 60,000 - 90,000€ [1]. Having this wide range of income in the area will allow this pilot

to evaluate the role of income in quality of life of the patients.

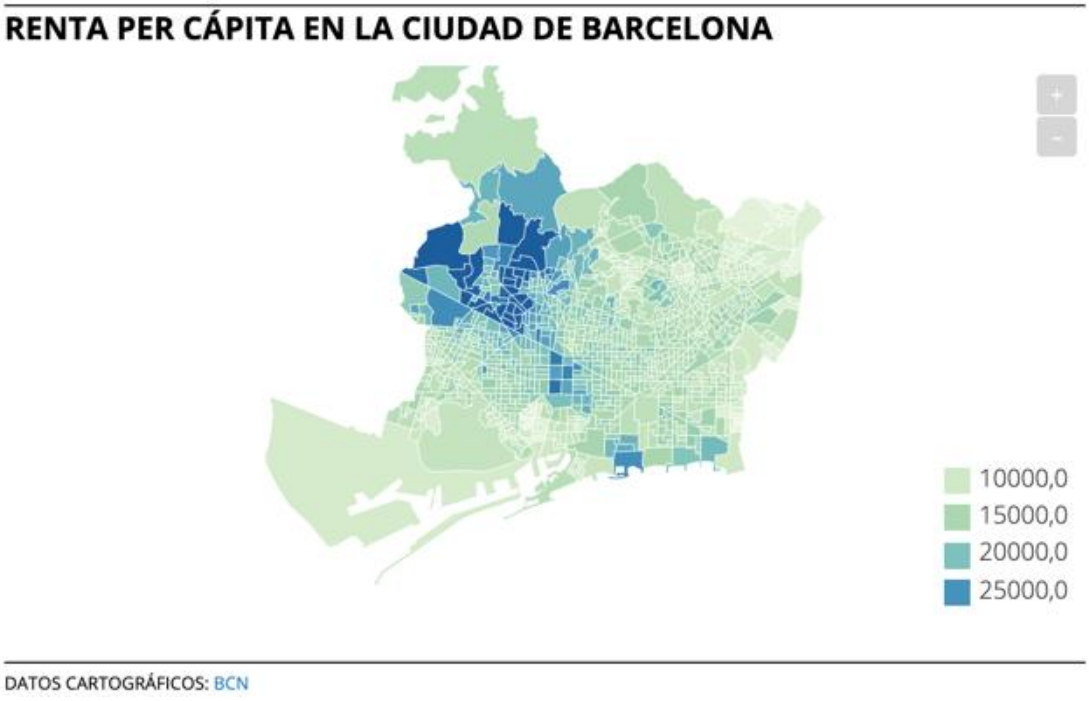

Figure 1. Map of Barcelona with the GDP of each neighbourhood.

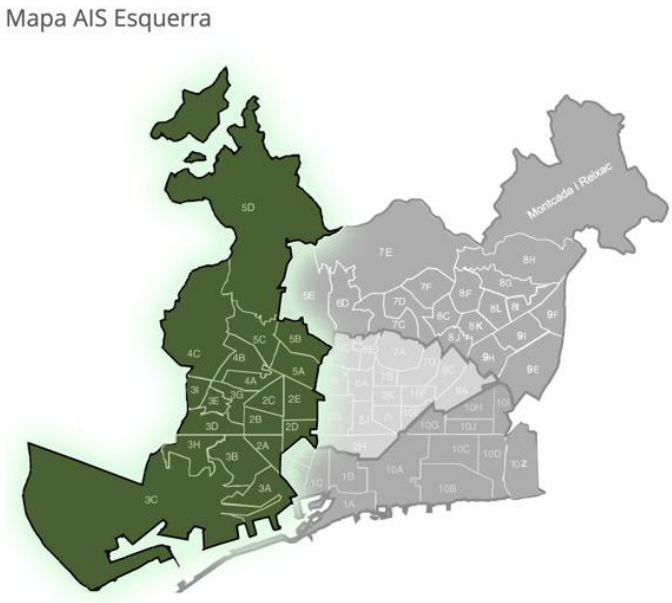

Figure 2.Highlighted on the map of Barcelona is the geographical area that is covered by Hospital Clinic and its associated primary health care centres.

2.2.1.2 Study procedure, data collection and use of ASCAPE

The Barcelona pilot will include ICT mainly using two methods: ASCAPE and the Xemio app. The combination of ASCAPE with the Xemio app demonstrates one of the originally envisaged ways for ASCAPE interacting, indirectly, with patients. However, the Xemio app goes beyond offering the convenience of data collection and the receipt of doctor-approved messages to patients on a mobile device. Indeed, Xemio has been designed to address the needs of breast cancer patients by including them in a support network in the area of Barcelona with actions such as meetings and live webinars with cancer experts who can answer the patients' questions. The Barcelona pilot aims to not only help collect data and evaluate ASCAPE with Xemio, but also Xemio in isolation (during the period before the First ASCAPE Prototype becomes operational). This results in additional complexities in the design of the Barcelona pilot, but also additional benefits as it will not only focus on ASCAPE proper, but also in a compatible and complementary ICT innovation for the support of breast cancer patients.

With regards to ASCAPE, the Barcelona pilot is slightly different to the other clinic-led pilots. All patients of this pilot will be recruited before the ASCAPE First Prototype is operational, therefore there is group of patients that will only contribute data. The other group of patients will finish their participation after the ASCAPE First Prototype is ready so they will receive the benefits of ASCAPE as ASCAPE AI results will be available to their doctors.

The Xemio app will be compatible with Android and iOS. It will also transfer data to ONTO-CR, which will house all the data in this study and will be the platform connected to ASCAPE. Participants using it, will be able to answer the questionnaires using Xemio directly and it will also provide daily steps and the quality of life issues that the participants experience and log in the app. Prior to recruitment, the study team will develop an information sheet with information on how to use the app. Finally, we will be in close contact with our physician partners to ensure that they are informed and aware of the study.

Upon approval by the Ethics Committee of the HCB, recruitment will start. The treating physician will ask the potential participants if they are interested in participating in the study. If they are, the field researcher will schedule a visit with the participant to review and sign the informed consent. Once a participant is enrolled, she will be randomised into one of the two groups: intervention with Xemio or control group without Xemio. The statistical analysis will be subcontracted to an outside group and they will manage the randomisation of groups in the Barcelona pilot. Randomization will be done using the PROC PLAN by SAS following a proportion 1:1 in blocks of 6. The randomization list will not be blinded, however the block order will be. Sample size calculation will not be performed because there are not enough patients to obtain the necessary data. If the participant is randomised to the control group, the field researcher will provide the participant with the questionnaires to be filled out. If the participant is randomised to the intervention group, the field researcher will provide the questionnaires to be completed by the participant, help

install the Xemio app to her phone and sign her up for the app. The field researcher will also train the participant on how to use the app and will give her the information sheet in case that she has questions later on.

As outlined earlier, the intervention will take place for 12 months. Every month the participants in the group using the Xemio app will receive a notification thanking them for their participation. The app also provides a calendar of events organised by breast cancer support groups that the participants can sign up for. Every 3 months all participants will receive the quality of life questionnaires as well as other questionnaires electronically (email or app) and will answer them on their own.

As soon as ASCAPE predictions and suggested interventions are available, the doctor will meet with the patient during a visit, in which he/she will see the predictions made by ASCAPE, decide if the patient should be made aware of those predictions, and whether to accept ASCAPE's intervention recommendations or to make different recommendations on the basis of their expertise and knowledge of the patient.

At the end of the intervention the control group will complete the quality of life questionnaires. The intervention group using the Xemio app will also complete the quality of life questionnaires and will participate in a focus group to collect the participant's impressions on using the Xemio app.

2.2.1.3 Quality of life measures

The following questionnaires will be used to capture different aspects on patients' quality of life:

- EORTC QLQ C30: at baseline and every three months
- Hospital Anxiety and Depression Scale (HADS): at baseline and every three months
- Loneliness questionnaire: at baseline and every three months.
- Health determinants such as smoking status, social status: baseline and twelve months.

2.2.2 **Orebro/Uppsala Breast Cancer Study**

2.2.2.1 Study design and population

The Örebro pilot has been designed as a prospective, single-arm, longitudinal study. Patients with early breast cancer planned for curative treatment with surgery with or without oncological treatment at the Örebro University Hospital or Uppsala University Hospital, respectively will be eligible for study inclusion.

The patients must fulfil all the following inclusion criteria:

1. Histologically proven breast cancer (any subtype).
2. No clinical evidence of metastatic disease.
3. Able for curative treatment with surgery with or without oncological treatment (endocrine therapy, chemotherapy, radiotherapy as neoadjuvant or adjuvant).
4. No prior malignant tumor during the previous 5 years, except for in situ carcinomas of the cervix or basal or squamous cell carcinomas of the skin adequately treated.
5. Signed informed consent before study entry.
6. Ability to utilise smartphone, apps, and wearables.

The patients will be excluded from the study if at least one of the following exclusion criteria is present:

1. Patients with metastatic disease not able for curative treatment.
2. Patients with known history of allergy to the wearable material.
3. Inability to give informed consent.

The healthcare pathway for breast cancer patients who will be eligible for the study is the same irrespectively the treatment approach. All newly diagnosed breast cancer patients are discussed on the weekly multidisciplinary team meeting. All patients are, thereafter, referred to the Department of Oncology either before (when preoperative treatment is recommended or in case of primary metastatic disease) or after surgery. The follow-up of breast cancer patients after treatment is performed by the Department of Oncology, mainly through dedicated oncology nurses. ASCAPE will, therefore, be integrated into the current follow-up practice of breast cancer patients as an additional tool for oncologists and oncology nurses.

As the study does not include a control arm, no formal sample size calculation has been performed. Considering the catchment area of the two healthcare regions (Uppsala Region län and Örebro Region län), the number of newly diagnosed breast cancer each year, and the expected inclusion rate, approximately 150 - 200 women are expected to be included during one year.

2.2.2.2 Study procedure and data collection

All potentially eligible patients for the study will be identified through the multidisciplinary team meetings. The patients will be informed about the possibility to participate to the study at their first visit at the Department of Oncology in Örebro University Hospital or Uppsala University Hospital. Informed consent will be obtained at the first visit.

At baseline (after informed consent), the following data will be collected: patients' characteristics, tumour characteristics, treatment-related information, questionnaire about lifestyle, baseline QoL measurements. Each included patient will also receive a wearable with ability for continuous measurements of heart rate, activity parameters, and sleep parameters. Each patient will have the wearable during the whole study period of 12 months.

The follow-up will be performed every 3 months using quality of life questionnaires for up to 12 months. An oncologist will be responsible for the follow-up with possibility to delegate the responsibility to a dedicated oncology nurse.

During the study follow-up (12 months for each patient), any intervention aimed to improve a measured quality of life aspect will be captured and additional follow-up (within 1 to 4 weeks depending on the nature of intervention) will be organised to assess the efficacy of the proposed intervention.

To capture patients' experience with ASCAPE-based follow-up, we will use a mixed-methods analysis (initial quantitative approach followed by qualitative methods to interpret the initial quantitative results). The focus of mixed-methods analysis will be on patients' satisfaction with ASCAPE-based follow-up, barriers and facilitators of using wearables during follow-up, and motivation for following interventions based on ASCAPE-based follow-up.

Similar methodology will be used for clinicians and oncology nurses to investigate barriers to recruitment and implementation of ASCAPE-based follow-up and how feasible and useful is to incorporate AI-based follow-up in current clinical practice.

All data will be collected to an electronic data capture system (Smart-Trial) using custom electronic case report forms. An encrypted link that will lead to the questionnaires will be sent by e-mail to the included patients through Smart-Trial both in pre-specified time points (at baseline and then every three months) but also after each proposed intervention for improving a quality of life measure (the time and frequency of questions to assess the efficacy of an intervention will be decided by the treating oncologist).

2.2.2.3 Quality of life measures

The following questionnaires will be used to capture different aspects on patients' quality of life:

- BREAST-Q (mastectomy or breast conserving therapy) at baseline
- BREAST-Q postoperative: every 3 months after surgery
- EORTC QLQ C30: at baseline and every 3 months
- EORTC QLQ-BR45: at baseline and every 3 months
- Hospital Anxiety and Depression Scale (HADS): at baseline and every three months
- Disability of Arm, Shoulder and Hand questionnaire (DASH): every three months

All questionnaires are extensively validated and are recommended by the Breast Cancer Evaluation Database to Guide Effectiveness (EDGE) Task Force about the clinical measures of health-related quality of life [2, 3].

2.2.3 CareAcross Breast Cancer Study

2.2.3.1 Study design and population

The CareAcross pilot for Breast Cancer will be based on the CareAcross personalised platforms for breast cancer patients. More specifically, individual patients will be informed and invited to use a version of the CareAcross service tailored to the ASCAPE infrastructure.

The inclusion criterion is Breast Cancer diagnosis (self-reported).

Given the nature of the platform, and that the patients self-register on it, and use it on their own, we assume that they are willing and able to use digital means to provide information and consume corresponding material.

CareAcross expect about 75-100 breast cancer patients to enrol.

2.2.3.2 Study procedure and data collection

Upon electronically receiving and electronically agreeing with the Informed Consent form, patients will be redirected to use the ASCAPE-specific CareAcross version of the online Breast Cancer service (a dedicated web platform accessible by any device with internet capability). Data will be collected directly from patients, who will be able to report upon specific aspects of their QoL, on a periodic basis.

Beyond the direct data input from patients, some patients will receive the designated wearable device, which will also contribute data on their behalf and based on their consent.

The nature of the CareAcross pilot is distinct from the clinic-led pilots, as the treating oncologist is not participating. It exhibits an important characteristic of the ASCAPE value proposition: ASCAPE democratises access to AI and Big Data related innovations in cancer patients' care, even if their hospital and/or clinicians are not participating. This can enable remote and long-term improvement of QoL in a scalable manner. At the same time, ASCAPE is not meant to provide direct advice or provide predictions to patients; a qualified medical expert is required by ASCAPE's design to both assess ASCAPE AI results and equally importantly which of those should be shared with patients, when and in what manner.

For the purposes of this pilot and within the context of the ASCAPE project, a clinician at CareAcross will be performing the corresponding task that the "treating physician" would perform. Therefore, the medical expert at CareAcross will perform the corresponding monitoring and identify the most relevant intervention, which will then be communicated to each individual patient, with the necessary wording. This wording will communicate, among others, that patients should understand that this is an experimental service, that it does not constitute medical advice, that it does not

replace their clinician's medical advice, that the ASCAPE project, CareAcross, and any other partner are not responsible or liable for these AI-driven interventions, etc.

The modality and frequency of this monitoring will be governed by operational and medical decisions, among others.

2.2.3.3 Quality of life measures

The data collected from breast cancer patients will reflect the following aspects of their QoL:

- Anxiety
- Depression
- Joint Pain
- Fatigue
- Neurotoxicity
- Hot Flushes
- Weight gain.

More specifically, the Anxiety and Depression aspects of QoL will be captured using the GAD7 & PHQ9 questionnaire, respectively. Weight gain will be captured by the actual weight being recorded. Each of the rest of the aspects will be captured through a single Boolean or Likert-scale input.

2.3 Prostate cancer studies

Three of the four pilot will engage in studies focusing on prostate cancer patients' QoL, the Athens pilot led by NKUA, the Orebro pilot led by Orebro University and the CareAcross pilot led by CareAcross.

2.3.1 **Athens prostate cancer study**

2.3.1.1 Study design and population

The Athens pilot is led by the Second Urology Department of Sismanoglio General Hospital. The trial is designed as a prospective, single-arm study. Patients with localised prostate cancer planned for curative treatment by surgically removing prostate gland along with seminal vesicles and ejaculatory ducts and also draining lymph nodes (from now on called radical prostatectomy) at Urology Department will be eligible for study inclusion.

The patients must fulfil all the following inclusion criteria:

1. Males aged > 18 y/o

2. Histologically proven prostate cancer (any subtype).

3. No clinical evidence of metastatic disease.

4. Able for curative treatment with surgery with or without oncological treatment (hormonal therapy, chemotherapy, radiotherapy as neoadjuvant or adjuvant).

5. No prior malignant tumour during the previous 5 years, except for basal or squamous cell carcinomas of the skin adequately treated.

6. Signed informed consent before study entry.

7. Ability to utilise smartphone, apps, and wearable devices.

The patients will be excluded from the study if at least one of the following exclusion criteria is present:

1. Patients with metastatic disease not able for curative treatment.

2. Patients with known history of allergy to the wearable material.

3. Inability to give informed consent.

4. Inability/ no access to use smartphones, applications, wearable devices or internet services.

After cancer diagnosis is confirmed with a prostate biopsy, the treating urologist advises the patient about potential therapeutic options according to patient malignancy, comorbidities, age, as well as his needs and expectations. The treating physician also informs the patient and his relatives about the pros and cons of each method. Finally, the patient after taking into consideration every aspect, makes his decision on which therapeutic pathway to follow. This "first contact" between patient-urologist, takes place into either the private office or the hospital outpatient appointments, depending always on patient's choice.

In case of surgery, which is the most common type of treatment performed in our department for localised cancer, the patient is admitted to Sismanogleio Hospital and interviewed regarding his social, medical and surgical history. The treating physician, along with a urology resident, informs the patient regarding the surgery once again. The patient is given a specific consent form which describes the procedure and complications and is asked to sign if he agrees to proceed with the surgery. Every potential question that may arise is answered both by the treating urologist and the residents who are on duty in our department.

In terms of participation in ASCAPE project, the patient will be fully informed by the treating urologist, during the initial appointment. In case the patient is willing to participate and fulfils the inclusion criteria, once admitted he will be given an extra consent form and the relative form for GDPR in order to be signed. Similarly, the treating physician and the residents will be constantly available for the patient in

case queries arise.

After patient receives his treatment, he will be followed at specific intervals. During this period of time, he will be able to contact a specific nurse and a resident of our department by phone call or e-mail, whatever he prefers in case of any concern or question regarding the filling of questionnaires or the use of the wearables.

As the study includes no control group, no formal sample size calculation has been performed. The expected number of patients during the anticipated period of recruitment after considering necessary period of time for active follow-up and the ability of patients, usually older than 65 y/o (the usual age of diagnosing prostate cancer is 50).

2.3.1.2 Study procedure and data collection

All potentially eligible patients for the study will be identified during initial appointment with their treating urologist. The patients will be informed about the possibility to participate in the study and in case of agreement and fulfilment of all predefined inclusion criteria, they will also be provided with a wearable device for use before surgery in order to record their baseline habits. The period of use before surgery is estimated to be 14 days. Informed consent will be obtained upon agreement to participate and after admission to the Urology Department of Sismanoglio hospital.

At baseline (after informed consent), the following data will be collected: patients' characteristics, tumour characteristics, treatment-related information, questionnaire about lifestyle, baseline quality of life measurements. Each patient participating in the study will be given the chosen wearable device and receive help in setting it up before leaving the hospital; they will keep it and be advised to wear it during the 12 months of their participation.

The follow-up will be performed every 3 months using quality of life questionnaires for up to 12 months. The treating Urologist along with a resident of the clinic and a dedicated nurse will be responsible for the follow-up. In case of delays in completion of questionnaires or upload of data received from the wearable device, the dedicated nurse will notify patient, in a polite manner, using e-mail or phone-calls, whatever is easier for the patient.

During the study follow-up (12 months for each patient), any intervention aimed to improve a measured quality of life aspect will be captured and additional follow-up (within 1 to 4 weeks depending on the nature of intervention) will be organised to assess the efficacy of the proposed intervention. Any intervention proposed will be decided after advising the treating urologist and explaining to the patient the pros and cons according to current higher level of evidence-based standard of care. The patient will decide which therapeutic option fits best for him. The role of ASCAPE platform will be to offer predictions and suggestions for the specific needs of every patient at an individual level, based on AI-platform training on similar cases. An advantage of ASCAPE platform is that it will help in applying knowledge gained from

hospitals/ centres that may specialise in offering care to prostate cancer patients, to a patient that may reside in remote places without access to such care, but even specialists will benefit from having access to AI results based on models created with input from colleagues around the world, or in the case of the ASCAPE prostate cancer studies from two other pilot sites.

To capture patient experience with ASCAPE-based follow-ups, we will use a mixed-methods analysis (initial quantitative approach followed by qualitative methods to interpret the initial quantitative results). The focus of mixed-methods analysis will be on patients' satisfaction with ASCAPE-based follow-up, barriers and facilitators of using wearables during follow-up, and motivation for following interventions based on ASCAPE-based follow-up.

Similar methodology will be used for clinicians and nurses to investigate barriers to recruitment and implementation of ASCAPE-based follow-up and how feasible and useful is to incorporate AI-based follow-up in current clinical practice.

Patient data in Urology Department of Sismanoglio Hospital are collected on paper and stored in a locked area with limited access only to responsible personnel. ASCAPE will be fully integrated with a system currently being implemented to facilitate the electronic capture of patient data. This is an update to the systems currently used in the Urology Department and potentially for the whole hospital later. A website will also be designed in order to be used from patient for questionnaires completion in specific time intervals every 3 months according to ASCAPE protocol. Each patient will determine his specific code, which will not be made public to the staff of ASCAPE. The first time of questionnaire completion, based on preoperative data depicting the QoL before radical prostatectomy, the staff of ASCAPE will explain to patients how the questionnaires are completed and aid them in case questions arise, without patronizing them. In case a QoL issue arises and an intervention is decided, then the relevant questionnaire may need to be completed in an altered time interval based on treating physician decision, in order to assess efficacy of the treatment.

2.3.1.3 Quality of life measures

The following questionnaires will be used to capture different aspects on patients' quality of life:

- EORTC QLQ C30: at baseline and every 3 months
- EORTC QLQ-PR25: at baseline and every 3 months
- Hospital Anxiety and Depression Scale (HADS): at baseline and every three months
- International Index of Erectile Function (IIEF): at baseline and every three months.

All questionnaires are extensively validated and are recommended by the Prostate Cancer Evaluation Database to Guide Effectiveness (EDGE) Task Force about the

clinical measures of health-related quality of life [4, 5].

2.3.2 Orebro prostate cancer study

2.3.2.1 Study design and population

The Örebro pilot has been designed as a prospective, single-arm, longitudinal study. Patients with localised prostate cancer planned for curative treatment with radiotherapy at the Örebro University Hospital will be eligible for study inclusion.

The patients must fulfil all the following inclusion criteria:

1. Histologically proven prostate cancer.
2. No clinical evidence of metastatic disease.
3. Able for curative treatment with radiotherapy irrespectively the type of radiotherapy (external, brachytherapy, combination). Patients with postoperative radiotherapy will also be included.
4. No prior malignant tumour during the previous 5 years, except for in situ carcinomas of the cervix or basal or squamous cell carcinomas of the skin adequately treated.
5. Signed informed consent before study entry.
6. Ability to utilise smartphone, apps, and wearables.

The patients will be excluded from the study if at least one of the following exclusion criteria is present:

1. Patients with metastatic disease not able for curative treatment.
2. Patients with known history of allergy to the wearable material.
3. Inability to give informed consent.

All newly diagnosed prostate cancer patients are discussed on the weekly multidisciplinary team meeting. All patients eligible for curative radiotherapy are, thereafter, referred to the Department of Oncology. The follow-up of prostate cancer patients after treatment is performed by the Department of Oncology, mainly through dedicated oncology nurses. The ASCAPE will, therefore, be integrated into the current follow-up practice of prostate cancer patients as an additional tool for oncologists and oncology nurses.

As the study does not include a control group, no formal sample size calculation has been performed. Considering the catchment area of the Örebro healthcare region, the number of newly diagnosed prostate cancer each year, and the expected inclusion rate, approximately 100 - 150 men are expected to be included for one year.

2.3.2.2 Study procedure and data collection

All potentially eligible patients for the study will be identified through the multidisciplinary team meetings. The patients will be informed about the possibility to participate to the study at their first visit at the Department of Oncology in Örebro University Hospital. Informed consent will be obtained at the first visit.

At baseline (after informed consent), the following data will be collected: patients' characteristics, tumour characteristics, treatment-related information, questionnaire about lifestyle, baseline quality of life measurements. Each included patient will also receive a wearable with ability for continuous measurements of heart rate, activity parameters, and sleep parameters. Each patient will have the wearable during the whole study period of 12 months.

The follow-up will be performed every 3 months using quality of life questionnaires for up to 12 months. An oncologist will be responsible for the follow-up with possibility to delegate the responsibility to a dedicated oncology nurse.

During the study follow-up (12 months for each patient), any intervention aimed to improve a measured quality of life aspect will be captured and additional follow-up (within 1 to 4 weeks depending on the nature of intervention) will be organised to assess the efficacy of the proposed intervention.

To capture patients' experience with ASCAPE-based follow-up, we will use a mixed-methods analysis (initial quantitative approach followed by qualitative methods to interpret the initial quantitative results). The focus of mixed-methods analysis will be on patients' satisfaction with ASCAPE-based follow-up, barriers and facilitators of using wearables during follow-up, and motivation for following interventions based on ASCAPE-based follow-up.

Similar methodology will be used for clinicians and oncology nurses to investigate barriers to recruitment and implementation of ASCAPE-based follow-up and how feasible and useful is to incorporate AI-based follow-up in current clinical practice.

All data will be collected to an electronic data capture system (Smart-Trial) using custom electronic case report forms. An encrypted link that will lead to the questionnaires will be sent by e-mail to the included patients through Smart-Trial both in pre-specified time points (at baseline and then every three months) but also after each proposed intervention for improving a quality of life measure (the time and frequency of questions to assess the efficacy of an intervention will be decided by the treating oncologist).

2.3.2.3 Quality of life measures

The following questionnaires will be used to capture different aspects on patients' quality of life:

- EORTC QLQ C30: at baseline and every 3 months
- EORTC QLQ-PR25: at baseline and every 3 months
- Hospital Anxiety and Depression Scale (HADS): at baseline and every three months
- International Index of Erectile Function (IIEF): at baseline and every three months.

All questionnaires are extensively validated and are recommended by the Prostate Cancer Evaluation Database to Guide Effectiveness (EDGE) Task Force about the clinical measures of health-related quality of life [6, 5].

2.3.3 CareAcross Prostate Cancer Study

2.3.3.1 Study design and population

The approach will be similar to that described in 2.2.3.1.

We expect about 50-75 prostate cancer patients to enrol.

2.3.3.2 Study procedure and data collection

The approach will be similar to that described in 2.2.3.2.

2.3.3.3 Quality of life measures

The data collected from prostate cancer patients will reflect the following aspects of their QoL:

1. Anxiety
2. Depression
3. Fatigue
4. Incontinence
5. Weight Changes
6. Sexual dysfunction
7. Hot Flushes

More specifically, the Anxiety and Depression aspects of QoL will be captured using the GAD7 & PHQ9 questionnaire, respectively. Weight gain will be captured by the actual weight being recorded. Each of the rest of the aspects will be captured through a single Boolean or Likert-scale input.

2.4 Per Pilot Experience Differentiations

Task 1.1 and Deliverable 1.1 delved into user-centric requirements analysis that resulted in use cases and related functional and non-functional requirements for an ASCAPE-powered healthcare information system. The ASCAPE vision is that such systems will be built on the basis of the ASCAPE framework and its components, but

may vary significantly in terms of user experience.

- User experience for doctors is largely dependent on the overall UI/UX design of the information system; for ASCAPE-provided functionality, the ASCAPE Dashboard provides a design proposal which the designers of the information system may or may not take under consideration.
- The experience of patients will largely be affected by their health status and the overall interaction with the healthcare provider; focusing strictly on ASCAPE-related aspects of their experience, the key differentiation points will be the ways they provide input to ASCAPE, directly (e.g. web forms completed by the patient, or by a healthcare professional) or indirectly (e.g. wearables), and any means of getting ASCAPE-provided output (e.g. doctor appointment, doctor-approved communication of ASCAPE results to the patient via electronic means).

Similar variations regarding the experience of doctors and patients are expected across 4 ASCAPE pilots.

2.4.1 Healthcare Provider Experience Differentiations

2.4.1.1 Overview and Rationale

The ASCAPE Dashboard Web Application is meant as a tool that demonstrates the capabilities ASCAPE has to offer, coupled with a concrete UI design proposal for making those capabilities accessible to doctors. It will be developed under the pilots work package, WP4. It will be used in the Orebro-led pilot along with an electronic data capture software and the current electronic medical record system and in the CareAcross pilot alongside the company's software system. It will not be used in the Athens and Barcelona pilots as its functionality will be fully integrated into the relevant corresponding IT Systems, thus showcasing the proposed means of making ASCAPE functionality in hospital systems across Europe.

- The case of the Orebro pilot is typical for a research project like ASCAPE; the introduction of a new system, the ASCAPE Dashboard, in the clinical practice of doctors participating in the Orebro University-led ASCAPE pilot will come with its difficulties but these will be manageable given the dedication and level of involvement with ASCAPE of the small group of doctors who will be actively participating in the Orebro University-lead pilot studies at the Department of Oncology at Örebro University Hospital (breast and prostate cancer) and the Department of Oncology at Uppsala University Hospital (breast cancer). ASCAPE being tested and proven useful during the pilot will be a prerequisite for any discussion of integrating ASCAPE functionality in the relevant information systems of the Departments of Oncology at the Hospitals.
- The CareAcross pilot is unique among the four ASCAPE pilots as CareAcross is not a medical healthcare provider having direct contact with patients but rather an SME focused on providing support to remote cancer patients

through education about their condition. CareAcross's service to the about 10,000 patients, is based on a (mobile-friendly) web-based platform that allows it to offer personalised informational content on the basis of questionnaires about patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) collected by the registered patients. The same communication channels will be used in the CareAcross ASCAPE pilot, but as ASCAPE is not designed to provide its predictions and intervention suggestions directly to patients, but to doctors monitoring their health status, the CareAcross ASCAPE pilot will diverge from business as usual, by adding not only ASCAPE functionality but also the element of CareAcross oncology experts deciding which intervention suggestions or other ASCAPE outputs are communicated to the patients enrolled in the CareAcross pilot. Given the small number of doctors and the fact that they will be dedicated to ASCAPE, the usability issues of having two systems to use, the CareAcross system and the ASCAPE Dashboard, are manageable. CareAcross expects to consolidate their exploitation plans for ASCAPE on the basis of the experience gained and the level of production readiness that will be achieved and decisions on future UI-level integration will depend on how CareAcross will make use of the ASCAPE project results to provide new services or enhance its current ones.

- The Barcelona pilot will involve, in addition to clinicians at HCB, doctors in a network of primary care centres in the metropolitan area of Barcelona. It is clear from the outset that any compromises in the user experience will significantly affect the ability of the Barcelona pilot to produce its desired results; doctors in this setting can be expected to only make use of ASCAPE functionality if it is readily available to them with little or no effort and with little or no prior training. For this reason, the Barcelona pilot will be one of the two pilots where full integration is planned to be achieved; the partners leading the Barcelona pilot and the partner leading the ASCAPE Dashboard will collaborate closely in order to provide a fully integrated experience, along the lines of what is hoped to be achieved by ASCAPE-powered healthcare provider information systems envisaged in D1.1.
- The Athens pilot is the second pilot in which full integration is planned to be achieved, albeit under significantly different conditions. The Information System for the Athens pilot is being designed and implemented at the moment and is expected to be largely influenced by the design of ASCAPE and of the Athens pilot. Whereas the Barcelona pilot will be a test of the ASCAPE design in upgrading an existing information system, the Athens pilot will be a test of its ability to inspire the design of modern AI-based user-friendly healthcare systems that provide services supporting cancer patients' QoL. This is also expected to be the result of tight collaboration of the involved parties and user experience is a high priority.

2.4.1.2 Use Cases

In D1.1 a small number of generic use cases was presented aiming to capture both interaction of the doctor (direct interaction) and of a patient (indirect interaction) with ASCAPE. These are repeated below for reference with additional notes highlighting pilot-specific differences. These pilot-specific notes do not point to exceptions to the requirements, nor they imply deviations from the design that is underway in T1.3; rather, they demonstrate how ASCAPE's requirements and design foresee variations in different settings.

Use Case HP.1 – Patient Visit

ID	HP.1
Name	Patient Visit
Description	This use case focuses on the interaction of the Doctor with the ASCAPE-powered Healthcare Provider Information System in the context of a Patient Visit.
Actors	The Doctor, The Patient
Preconditions	The Patient has entered the optional ASCAPE personal data collection scheme offered by their healthcare provider (See Use Case PT.1 below)
Trigger	Patient visit to the Doctor
Main path	1. The Patient enters the Doctor's Office 2. The Doctor finds the Patient's record in the ASCAPE-powered Healthcare Provider Information System. 3. The System logs the fact that the Doctor has visited the Patient's record 4. The Patient's record, thanks to ASCAPE-integration, offers features such as: a. A list of ASCAPE signals regarding the patient's QoL metrics (current and/or predicted) b. What-if Graphs of QoL metrics visualising the historic and/or predicted values of QoL metrics and the potential effect of proposed interventions; the default configuration when the Doctor opens the Patient's record presents a comparison of any ASCAPE-proposed interventions with the no-intervention case 5. The Doctor quickly peruses the QoL signals, by clicking on each one of them. There are two kinds of signals: a. Purely informative signals highlight a concern about the patient's quality of life either on the basis of current data or on the basis of a predicted trajectory of one or more indicators, where the prediction can be made high sufficiently high accuracy; upon the Doctor clicking on one of those signals the graphs for the relevant metrics appear, clearly distinguishing between observations up to that day and predicted values, as well as providing confidence levels for the latter

	b. QoL intervention suggestion signals, which appear when there is sufficiently high confidence that an intervention will improve the Patient's QoL, show information about why the specific intervention suggestion is made. 6. The Doctor, curious to see why another possible intervention was not recommended, uses the What-If Graphs manually to see ASCAPE's view on that intervention and then proceeds to obtain an explanation from ASCAPE about the basis on which it proposed the interventions that it did 7. The Doctor consults with the Patient and determines if there are any medical reasons, possibly not recorded in the patient's history, making any of the considered interventions inappropriate, updates the patient's history if necessary, and judges what the optimal course of action is appropriate for the Patient. 8. The Doctor discusses the recommended course of action and any alternatives with the Patient (including the wait-and-see/no-intervention option where appropriate) and shares ASCAPE predictions about each of option if they find them to be in accordance with their own medical opinion 9. The Doctor and the Patient agree on a course of action (which the Doctor records in the ASCAPE-powered IT system) and discuss what the Patient should do, what symptoms to look out and report to the Doctor (or a colleague of a particular specialisation where appropriate). and what expectations they should have 10. The visit and the agreed course of action are recorded by the System 11. The Doctor and the Patient renew their appointment either for after the regular monitoring interval or earlier if appropriate.
Postconditions	The visit, the Doctor's access to the Patient's record and the agreed course of action for the Patient are recorded.
Athens Pilot	• The Doctor will have access to ASCAPE-powered functionality directly through the Sismanoglio Prostate Cancer IT System; the ASCAPE-related UI elements will be based on the ASCAPE Dashboard's design • ASCAPE-powered functionality will be based on the prostate cancer models
Barcelona Pilot	• The Doctor will have access to ASCAPE-powered functionality directly through the Barcelona OntoCR Breast Cancer IT System; the ASCAPE-related UI elements will be based on the ASCAPE Dashboard's design • ASCAPE-powered functionality will be based on the breast

	cancer models
CareAcross Pilot	• The Doctor will have access to ASCAPE-powered functionality directly through the ASCAPE Dashboard • ASCAPE-powered functionality will be based on the breast and prostate cancer models
Orebro Pilot	• The Doctor will have access to ASCAPE-powered functionality directly through the ASCAPE Dashboard • ASCAPE-powered functionality will be based on the breast and prostate cancer models

Use Case HP.2 – Doctor Alert

ID	HP.2
Name	Doctor Alert
Description	This use case focuses on the ability of an ASCAPE-powered Healthcare Provider Information System to alert the Doctor about a developing health issue with the Patient
Actors	The Doctor, The Patient
Preconditions	The Patient has entered the optional ASCAPE personal data collection scheme offered by their healthcare provider (See Use Case PT.1 below)
Trigger	Input from Patient or authorised health-monitoring device
Main path	1. Input from Patient or authorised health-monitoring device reaches the ASCAPE Platform, is processed and is converted into an ASCAPE signal 2. An informative indicator prompts the Doctor to view the new task in the ASCAPE-powered Healthcare Provider Information System. The task description explains the issue identified by ASCAPE, the timestamp and the type of data that triggered the signal (e.g. wearable, self-reporting questionnaire or other) and contains a link to the patient's record 3. There the Doctor finds the complete list of ASCAPE signals regarding the patient's QoL, quickly examines them (e.g., system indicates different entries) and determines that it would be best to arrange a call with the Patient. 4. The Doctor uses the ASCAPE-provided functions of the Healthcare Provider Information System during the call with the patient, in a manner similar to what was described in Use Case HP.1 above.
Alternate path	
Postconditions	
Athens Pilot	• The Doctor will have access to ASCAPE alerts directly through the Sismanoglio Prostate Cancer IT System
Barcelona Pilot	• The Doctor will have access to ASCAPE alerts directly through ONTO-CR
CareAcross	• The Doctor will have access to ASCAPE alerts through the

Pilot	ASCAPE Dashboard to which they will have to log in; since this pilot lacks the element of the patient's physical presence, the alerts functionality will guide the Doctor in deciding which patients' health status they will need to review and will be central to his/her interaction with ASCAPE
Orebro Pilot	• The Doctor will have access to ASCAPE alerts through the ASCAPE Dashboard to which they will have to log in separately to other systems they will need to use in their daily routine

2.4.2 Patient Experience Differentiations

2.4.2.1 Overview and Rationale

In an ASCAPE-powered information system, patients are, ultimately, the beneficiaries of the support the ASCAPE AI can provide to their doctors in following their overall health and QoL status, but do not come in direct contact with the ASCAPE framework and its supporting infrastructure. After informed consent, patients contribute relevant data to ASCAPE for predictions and personalized suggestions for interventions. The supported modalities of data collection are determined by the healthcare provider, which is another area of experience differentiation.

- Three pilots will use a wearable (Athens, CareAcross, Orebro)
- One will use a specially designed mobile app (Barcelona)
- All four will use, in all or some cases, web-sites where patients can fill in questionnaires (Athens, Barcelona, CareAcross, Orebro)
- Three may involve questionnaires being filled with the help of staff during the initial completion at baseline and beyond, where needed (Athens, Orebro, Barcelona)

2.4.2.2 Use Cases

Use Case PT.1 – Patient interaction with ASCAPE

ID	PT.1
Name	Patient interaction with ASCAPE
Description	This use case focuses on how a Patient interacts with an ASCAPE-powered Healthcare Provider Information System. Unlike other use cases which concentrate on a short event, this use case covers the duration of a patient's interaction with ASCAPE. (The patient's interactions with ASCAPE listed herein are indicative, are not mandated nor dictated by the ASCAPE

	framework and may different from Healthcare Provider to Healthcare Provider and from patient to patient)
Actors	The Patient, The Doctor
Preconditions	
Trigger	
Main path	1. The Patient enters the optional ASCAPE personal data collection scheme offered by their healthcare provider and supported by the ASCAPE Cloud and the ASCAPE-powered Healthcare Provider Information System: a. The Patient signs the ASCAPE Data Processing Consent Form or a form with the same legal effect provided by their healthcare provider in order to benefit from the ASCAPE services b. The patient provides access to data from a wearable provided by their healthcare provider (optional, but without this the patient's ASCAPE predictions and intervention suggestions will not benefit from data from a wearable device) c. The patient downloads a mobile application provided by their healthcare provider and grants it access to send data (e.g. short QoL questionnaires) to the healthcare provider's ASCAPE-powered Information System (optional, but without this the patient's ASCAPE predictions and intervention suggestions will not benefit from data from the Healthcare Provider's mobile app) 2. The Patient provides data to the ASCAPE-powered Healthcare Provider Information System either directly (e.g. via questionnaires they fill in) or indirectly (e.g. via the wearable device) and so does the healthcare provider (e.g. information about what QoL improvement intervention they recommended to the patient) a. The wearable device provides data on the Patient's physical activity and sleep quality (alternatively it could provide additional information, such as blood oxygenation, or not be used at all) b. The mobile app may be used for reporting regularly side effects, daily steps, level of participation in social meetings related to breast cancer. And the participants are able to fill QoL mini-surveys or receive links for filling more detailed QoL questionnaires c. The Patient is also asked to fill in more detailed QoL questionnaires on tablets given to them during their visits to the Healthcare Provider's premises as well as from the comfort of their home by using the Healthcare Provider's website (a number of further alternatives could be seen in practice, including having a member of staff ask the questions and record the answers either on

	the Healthcare Provider's premises or over the phone) d. The Patient's personal information, medical history, cancer treatment information, doctor visits and lab results are also entered into the ASCAPE-powered Healthcare Provider Information System Note: None of these data are transmitted by the ASCAPE-powered Healthcare Provider Information System to the ASCAPE Cloud, but the knowledge they contain is made available via advanced technical means in accordance with the Healthcare Provider's policies as reflected by the relevant system settings (see use case ADM-1). 3. The doctor can use AI-provided functionality provided by the ASCAPE-powered Healthcare Provider Information System based on the above data to consult the patient, suggest appropriate interventions, or adjust the follow-up schedule. See use cases HP-1 and HP-2. 4. The Patient's monitoring ends. Patient data stored on the ASCAPE-powered Healthcare Provider Information System may be subject to (partial) deletion in accordance with the Healthcare Provider's and the exact terms of the Patient's consent form (see Step 1.a). The knowledge obtained from the Patient's data remains in the ASCAPE Cloud for the benefit of other patients. Steps 2 and 3 may be repeated and as data collection and medical follow-ups are part of a continuous process.
Alternate path	
Postconditions	
Athens Pilot	Step 1: • Patients participating in the Athens Prostate Cancer Study will be included after signing the relevant consent forms, after detailed discussion with the treating physician and ASCAPE members. • If the patient agrees to use the ASCAPE-provided wearable, the necessary registration will take place with the help of ASCAPE members. A list of e-mail addresses and phone numbers will be provided to the patients and their relatives in case any question arises regarding wearable use. Step 2: • Patients will fill in online questionnaires; those with the ASCAPE-provided wearable will be providing data through the mobile phone tethered to the wearable Step 3: • Doctors will advise patients and/or communicate any ASCAPE results they judge appropriate when the patients visit them at Sismanoglio or during follow-up visits. Step 4:

	• After ASCAPE, patient data will remain in the hospital database in order to be used for future patient follow-up and standard of care.
Barcelona Pilot	Step 1: • Patients participating in the Barcelona Breast Cancer Study offer their consent by signing the informed consent document offered and explained by the field researcher. • If the patient is selected and agrees to use the Xemio app , the necessary registration will take place with the help of the field researcher. By signing the informed consent form the participant will be agreeing to share the data collected in the app with ASCAPE via ONTO-CR. Step 2: • Patients fill in questionnaires either on Xemio (if they have agreed to use it) or on a secure web form the URL of which they will have received by email (otherwise); those who use Xemio, will also have their Xemio-tracked activity data sent to OntoCR, the central IT system of the Barcelona pilot (and from there to the Barcelona ASCAPE Edge Node) Step 3: • Doctors at the Clinic hospital or the participating Barcelona Primary Care Centres will advise patients and/or communicate any ASCAPE results they judge appropriate when the patients visit them in the clinic. Step 4: • After ASCAPE, patient data will remain stored in ONTO-CR in order to be used for future patient follow-up and standard of care.
CareAcross Pilot	Step 1: • Patients participating in the CareAcross pilot offer their consent online; new patients do so during registration if they wish to participate in the ASCAPE pilots, selected-to-participate previously registered patients in a special online consent form Step 2: • Patients will fill in online questionnaires on CareAcross's site; those with the ASCAPE-provided wearable will be providing data through the mobile phone tethered to the wearable Step 3: • CareAcross registered patients will not be in direct contact with doctors reviewing their health status with ASCAPE's help and sending them personalised health advice and any ASCAPE results they see fit to share with them; communication will be via electronic means and will not aim to substitute the contact patients need to have with their

	doctors, as per CareAcross' model which is to facilitate, rather than replace such contact. Step 4: • After ASCAPE, patient data will be transformed to be used by the "traditional" CareAcross systems, as per the consent forms already in place. This way, it will be possible for patients to receive the "traditional" CareAcross functionality after ASCAPE, if desired.
Orebro Pilot	Step 1: • Patients participating in the Orebro Prostate Cancer Study offer their consent after they receive information about the scope of ASCAPE from their treating physician on their regular visit at the Department of Oncology, Örebro University Hospital. • Patients participating in the Orebro/Uppsala Breast Cancer Study offer their consent after they receive information about the scope of ASCAPE from their treating physician on their regular visit at the Departments of Oncology, Örebro University Hospital or Uppsala University Hospital. Step 2: • Patients will fill in questionnaires via email or directly during their visit with the help of dedicated oncology nurses; those with the ASCAPE-provided wearable will be providing data through the mobile phone tethered to the wearable Step 3: • Doctors will communicate any ASCAPE results and advise the patients during their regular follow-up visits at the Department or over the phone if the doctor and the patient has agreed upon this type of communication. Step 4: • After ASCAPE, patient data with impact on patients' follow-up and future care, as assessed by the treating physician, will be retained in the current electronic medical record system for future use.

2.5 ASCAPE Pilots Wearable

The terms "wearables" and "wearable devices" refer to smart electronic devices, typically with an array of sensors, which can comfortably be worn on the body or incorporated into clothing. Of particular interest for the healthcare sector are the fitness/activity trackers or wristbands, as well as smartwatches, albeit focusing mainly on their fitness/activity monitoring functionalities. Such devices, provide compelling opportunities and are extensively impacting the healthcare offerings as they enable the potential for non-invasive, constantly vigilant, and low-cost monitoring of the individual's condition and can not only help improve personal

health and well-being [7, 8], exactly as ASCAPE aims to do. For this reason, three of the pilots (Athens, CareAcross, and Orebro) include input from wearables in their pilot designs.

2.5.1 Selected Wearable

At the early stages of the wearable's selection process, it was decided that a single common wearable device would be used for all three pilots using wearables. This would ensure a common codebase for interoperability with the devices and a shared experience throughout the three pilots. The wearable device that was selected is Fitbit Inspire HR. It was selected on the basis of the process outlined in Section 2.5.3 which in turn was based on the requirements gathered via the process described in Section 2.5.2.

2.5.2 Requirements Gathering Process

The initial process to define the wearables requirements involved all clinical partners using the consensus decision making approach in each step.

At first, three distinct categories of requirements were considered: a) general data-specific, b) non-data specific requirements and finally c) data to be collected. In this process, the clinical pilots were asked to provide relevant requirements for general data-specific and non-data specific categories. After providing the potential requirements, a consensus was reached for nine general data-specific and nine general non-data specific requirements (see Table 7 in Annex I).

Regarding the data to be collected, we initially focused on 13 parameters that are able to be collected from commercially available wearables including steps, activity type, activity time, calories burned, heart rate, heart rate variability, oxygenation, sleep quality, sleep quantity, body temperature, gait analysis, and blood pressure. For each specific QoL issue (as defined in Section 3) the initial data collection requirements were identified (see Table 9 and Table 10 in Annex I), and in the next stage the clinical experts independently chose the parameters from wearables that are of interest for the specific QoL issue and prioritised them with A (highest priority) to E (lowest priority) and a consensus was reached among all chosen parameters and prioritisations (see Table 11 and Table 12 in Annex I).

After this process, the parameters from wearables that are of interest for the ASCAPE were selected based firstly on the total number of QoL issues that will use each specific parameter and secondly on the number of QoL issues with highest priority for each parameter. Using the above-mentioned process, the following parameters from wearables were chosen: steps, activity type, activity time, calories burned, heart rate, sleep quality, and sleep quantity (see Table 8 in Annex I).

A technical committee proceeded to evaluate the requirements and in collaboration with the pilot partners further refined them, eliminating ambiguities and introducing a

small number of additional data specific and non-data specific requirements that were deemed beneficial for the project:

- **The ability to export or extract data from the wearable device**: Several vendors are offering wearables devices that are operating in so-called "closed" ecosystems that are built around services, applications and other offerings, usually controlled by the same vendor, or in some cases by third-party collaborators of the particular vendor. In such ecosystems, the export and consequently exploitation or utilization of the generated data by other solutions such as the ASCAPE framework is not possible and, in many cases, prohibited or even blocked. Thus, such wearable solutions shall be excluded from the wearable selection process of the ASCAPE framework.
- **The cost of the wearable device**: In the wearables market, the various vendors are offering a large variety of wearable devices that incorporate numerous functionalities and services by integrating numerous sensor types and related hardware. In this sense, the cost of available wearable devices ranges from low-cost devices to devices with high cost. In accordance with the ASCAPE Description of Action and the collected requirements, wearables devices with high cost of purchase are not included in the wearable selection process of the ASCAPE framework. For this reason, vendors like Apple and their flagship wearable device Apple Watch and Samsung's flagship wearable devices Galaxy Watch and Galaxy Watch Active 2 are excluded from the selection process.
- **The previous experience of the ASCAPE partners**: The efficient and effective integration of a wearable solution to the overall ASCAPE framework requires several technical aspects to be taken into consideration. Within this context, the familiarity of the ASCAPE technical partners and previous experience with the respective vendor's integration offerings is considered as a crucial criterion during the wearable selection process. The rationale of this criterion is the reuse of as much as possible existing solutions from the partners towards the optimization of the effort consumption, as well as the implementation process.

2.5.3 Evaluation Process

The evaluation process begun with a list of devices from well-known manufacturers with acceptable costs; this included also some devices that ASCAPE partners had no prior experience with, however it ensured that all suitable devices are included in the evaluation process that would reveal the device that would better serve the ASCAPE pilots. In the initial steps of the evaluation, the assessment focused firstly on the ability of these devices to capture various types of data and secondly to their ability to provide them to ASCAPE. In the first step, the initial list included 12 candidate devices from which 2 devices were eliminated due to their failure to meet the complete list of parameters to be collected as defined in section 2.5.2 (see Table 13 in Annex I).

In the second step, the remaining 10 candidate devices were assessed against the three additional requirements set by the technical committee. From this assessment, only 2 devices were successfully evaluated as the rest of the candidate devices were eliminated due to limitations in their offered data export functionality and integration possibility with the ASCAPE or their limitation in the type and format of the data that can be exported (see Table 14 in Annex I).

In the final and third step, the two remaining devices, Fitbit Charge 4 and Fitbit Inspire HR were evaluated against the pilot partners' data-specific and non-data specific criteria (see Table 15 and Table 16 in Annex I). Both devices have equivalent features and functionalities with the exception of the built-in GPS sensor that is incorporated in the Fitbit Charge 4 device. However, the availability of this sensor in the context of the project is not deemed necessary and additionally it raises the cost of acquisition of the specific device. Hence, based on the results of the whole process the proposed wearable device is Fitbit Inspire HR.

2.6 Timeline of pilots within ASCAPE

According to the ASCAPE Workplan, the period for the preparation and execution of the pilots is M10 (October 2020) until M34 (October 2022) inclusive, with:

- M18 being the deadline for both for D4.1, the ASCAPE Dashboard, and D4.2, describing the trials setup,
- M24 being the deadline of both D3.2, the initial ASCAPE Integrated Prototype and D4.3, a report on the initial pilots execution and evaluation
- M34 being the deadline for D4.4, a report on the final execution and evaluation of the pilots,
- M36 being the deadline for D3.3, the final ASCAPE Integrated Prototype.

Not all pilots start at the same time and not all have the same goals and design but their specific timelines are in broad agreement with the overall workplan.

Pilot	Recruitment period	Participation Period	First Patient Data Input	Last Patient Data Input
Athens	M11-M			

While there are differences among them, the three clinic-lead pilots (Athens, Barcelona, Orebro) have a common double aim in their studies: to not only evaluate ASCAPE and offer opportunities for troubleshooting and improvement during that evaluation, but to collect data in accordance with their data specification even before an operational version of ASCAPE and any local integrations are available. The early data collection will help improve, evaluate internally, troubleshoot and optimise

the AI models, leading to a better ASCAPE at the start of the first pilot run and, by speeding up developments, a better ASCAPE overall.

There are two striking peculiarities in the clinic-lead pilot timelines, one concerning when the first patients are recruited in the relevant studies and one concerning when the patients participation ends. Both are the result of decisions for the benefit of ASCAPE and both are compatible with the workplan as well as in perfect alignment with the ASCAPE project's aims.

The first peculiarity is that while in accordance to the workplan the deadline for D4.2 describing the trials setup and D4.1 the ASCAPE Dashboard is in M18, all three clinic-lead studies (Athens, Barcelona and Orebro) start before that as can be seen in the descriptions of their timelines below. This is done in order to allow the early collection of data which will help train and evaluate the AI models at an earlier stage than it would have otherwise been possible and reflects the level of readiness and commitment of the respective pilot-leading partners. If this design decision was not taken by the three clinical partners but instead all pilots started collecting data exactly when an early version of the ASCAPE Integrated Prototype was put in operation, this early version of would have had to be released by the Consortium without the benefit of any internal evaluation, troubleshooting and optimisation based on any amount of ASCAPE prospective data and, likewise, would initially be producing results without the benefit of any amount of ASCAPE prospective data either. Given that retrospective data, while useful, do not cover important aspects of the ASCAPE models, not taking the decision to start collecting prospective data at an early point would have had a negative impact on the project, not only initially, but also in how far improvements could proceed until its end. The early setup of the three clinical pilots falls within the time limits of the relevant task, Task 4.2, and the collection of prospective data for the purposes of ensuring ASCAPE is ready to be used in the first pilot run falls within the task's description.

The second peculiarity in the clinic-lead pilot timelines concerns when patients' participation ends. While in accordance with the workplan M34 is the end month of Task 4.3. ASCAPE framework validation for Prostate Cancer and Task 4.4. ASCAPE framework validation for Breast Cancer as well as the deadline for D4.4 which reports on the final execution of the ASCAPE pilots and provides the relevant evaluation results, two clinic-lead pilots (Athens and Orebro) will keep collecting data from patients participating in their ASCAPE studies after M34 and indeed after the project has officially ended. This is a side-effect of the combination of the decision to keep each patient in the relevant ASCAPE studies for 12 months and the need to have extended recruitment periods in order to maximise each pilot's capability to sign up a sufficiently large number of patients in order to collect data that will assist in the training of the ASCAPE AI models and in the evaluation of ASCAPE's performance. From a technical perspective it is possible to keep patients on an ASCAPE study after the project's 3-year duration as the ASCAPE Cloud infrastructure will be available for at least 12 months after the project ends and the relevant ASCAPE Edge infrastructures will also be kept available by the clinics. Clearly data collected by patients recruited closer to the end of the recruitment

periods and concerning their follow ups closer to the end of the 12 month participation period will not be available for inclusion in the D4.4 which will be reporting on the execution and evaluation of the final pilot run, but the alternative would have been to have shorter recruitment periods, shorter participation periods, or both, and as a result have fewer patients participating, fewer data points per patient or both. Another alternative would have been to run studies with different participation periods, so that no data is collected in M34 or afterwards, which would complicate the study design and lead to less data being available overall. Data not considered in D4.4 will:

1. contribute in the continuous training of the ASCAPE models up until M36
2. contribute in delivering a better Final Integrated Prototype in M36,
3. contribute in the continuous training of the ASCAPE models after M36
4. be available for the benefit of patients as well as research activities after the end of the project.

The CareAcross pilot does not have the same complications that the clinic-lead pilots do as it is meant to start with ASCAPE operational. Recruitment for the CareAcross pilot will happen prior to that and once everything is ready on the technical side, patients that are enrolled in the CareAcross system, submit data to it about their QoL issues and receive advice on the basis of the rules-based CareAcross decision algorithm, will instead be presented with ASCAPE-CareAcross questionnaires and start receiving CareAcross-ASCAPE intervention suggestions. With leading university clinics

As in all clinic-lead pilots recruitment and participation in the studies begins without ASCAPE being available to doctors, there is a question about whether they are switched over to receiving ASCAPE benefits once ASCAPE is operational. If they are not switched over to benefiting from ASCAPE, they can be used as a kind of control group for the purposes of the health-economics evaluation which requires

will be data points that can be used for contrasting directly

Data collection after”

[The decision to allow collection of data in and after M34 does not affect the deadlines of D4.3, D3.3 or any other deliverables. Data collected up to M36 will be used for updating the models of M3.3, the final ASCAPE prototype, delivered in M36 and data collected afterwards will

- relevant data will be both research paper ncluded in models available for the evaluation as reported

This ensures that prospective data collection necessary to train the AI predictive models can start early in the project and the ASCAPE services base
d on available predictive models (M24, December 2021) will be applicable as part of the framework validation until M34 (October 2022) as planned. In order to allow for an evaluation of the ASCAPE services patients are divided into those only contributing to the development of the AI models and those benefitting from the ASCAPE services after M24. Each individual patient will enrol for a period of at most 12 months, which means that patients enrolling in M10-M24 will only contribute to the prospective data collection to develop the AI models and never be subject to the use of ASCAPE services. Patients enrolling from M25 on benefit from the ASCAPE services and contribute to the improvement of the AI models as part of the continuous learning. The patients enrolling in M10-M24 will be the control group and those enrolling in M25 and later will form the intervention group for ASCAPE and as necessary for the health economics evaluation (see Section 4.4).

3 Pilot Data and QoL Data Determinants

3.1 Retrospective and Prospective Datasets

3.1.1 Retrospective Medical Datasets and Open Data Examination

For breast cancer, five retrospective datasets from three partners are available to be used in the ASCAPE.

Barcelona pilot has a dataset of n = 203 early breast cancer patients treated with different treatment approaches and includes detailed information about patient- and tumour characteristics, treatment strategies, side effects, and QoL aspects gathered through validated questionnaires as EQ-5D. Of the 203 patients of the cohort, retrospective data will only be given from those patients that will not participate in the prospective pilot, therefore avoiding using retrospective and prospective data from the same participants.

CareAcross brings a dataset of n = 1,700 breast cancer patients with early or metastatic breast cancer treated with different treatment approaches. The dataset includes data about tumour characteristics, treatment strategies, and side effects based on patients' answers through an online platform.

Örebro pilot brings three retrospective breast cancer datasets. The Sörmland dataset is a research database including n = 232 early breast cancer patients treated with endocrine therapy after surgery. The aim of the Sörmland study was to investigate the incidence of joint pain due to endocrine therapy and its impact on several QoL issues as fatigue, insomnia, anxiety, and depression. Validated questionnaires (Brief Pain Inventory and Fibromyalgia Impact Questionnaire) were used for this purpose. The Uppsala dataset is a dataset aimed to investigate the impact of different surgical approaches on QoL. The dataset included n = 250 early breast patients who fulfilled validated questionnaires (Breast-Q and DASH) before and after surgery. Data on patient- and tumour characteristics as well as treatment strategies were also gathered. Finally, the BcBaSe is a research dataset based on national epidemiological databases in Sweden with data about patients with early breast cancer (n = 15,000) in terms of patient- and tumour characteristics, treatment strategies, prescribed medications, newly diagnosed diseases after breast cancer diagnosis, number and reasons of hospitalizations. This dataset does not include QoL issues captured through questionnaires but it is able to use proxy for some QoL issues of interest as anxiety, depression, insomnia, neurotoxicity where the prescription of a new medication with one of the above mentioned indications can be considered as a proxy for the presence of the specific condition.

Description of dataset	Number of Patients
Barcelona Breast Cancer Survivors Dataset	203
CareAcross Breast Cancer Patients with QoL Questionnaires Filled Out	1,700
Quality on life in patients treated with aromatase inhibitors as adjuvant therapy (Breast cancer – Sörmland)	232
Quality of life depending on type of surgery (Breast cancer – Uppsala)	250
BcBaSe (static database for research purposes based on National Quality Registry for breast cancer)	15,000

Table 1 Breast Cancer Retrospective Datasets

For prostate cancer, two retrospective datasets from two partners are available.

CareAcross brings a dataset of n = 250 prostate cancer patients with localised or metastatic prostate cancer treated with different treatment approaches. The dataset includes data about tumour- characteristics, treatment strategies, and side effects based on patients' answers through an online platform as well as data from validated questionnaires (GAD7 for anxiety, PHQ9 for depression).

Örebro prostate cancer dataset is a real-world database including patients (n=2340) with localised prostate cancer treated with radiotherapy with curative intention in the Örebro healthcare region and has as aim to be used as a tool to improve the follow-up of those patients. It includes data on patient- and tumour characteristics, treatment approaches, dosimetric parameters regarding radiotherapy, side effects based on direct questions or validated questionnaires (IPSS, IIEF-5) and QoL issues based on a validated questionnaire (LISAT-11).

Description of dataset	Number of Patients
CareAccross Prostate Cancer Patients with QoL Questionnaires Filled Out	250
Prostate cancer database in Örebro (Radiotherapy with curative intention)	232

Table 2 Prostate Cancer Retrospective Datasets

In order to identify potential open data sources of interest to the ASCAPE, an extensive review of datasets available on the web for breast or prostate cancer was

performed. The searching strategy yielded 53 datasets for breast cancer and 23 for prostate cancer. A critical review of the open datasets was done to investigate the relevance to the ASCAPE scope. None of the open datasets presented data on QoL issues and as a result, none of them is suitable for inclusion into ASCAPE.

QoL data from Eurostat will be reviewed in order to find potential sources of biases that can be present due to differences among countries or within other parameters (as gender, age, socio-economic status) and that would need to be considered for normalisation of ASCAPE datasets.

Although the potential impact of climate changes and variability on the human well-being and quality of life have been highlighted 20 years ago [9] a recent systematic review found limited evidence on this potential association and urged more research in this field [10]. It is, therefore, planned to utilise open environmental data (from European Environmental Agency) to investigate the potential impact of climate on the different QoL issues in cancer patients.

3.1.2 Data Collection and Prospective Datasets

During the four pilots, prospective data will be collected through a variety of sources depending on the type of data and the pilot design. The different sources include validated questionnaires, wearables, mobile app, and questions through web-based platforms.

3.1.2.1 Questionnaires

Validated questionnaires will be used to capture the QoL issues in which a validated questionnaire is available. The use of validated questionnaires is essential to minimise the risk for measurement bias, facilitate the open call process, and enable the generalisability of ASCAPE-based predictions and interventions.

In some QoL issues, different validated questionnaires among pilots will be used to capture the same QoL issue. Based on current evidence and clinical expertise within the ASCAPE consortium, the presence or absence of each QoL issue will be defined within each validated questionnaire and the definitions will be aligned.

Patients will be asked to answer the validated questionnaires chosen for each pilot study every three months. There are two exceptions with regards to the use of validated questionnaires among the pilots: the Barcelona and the CareAcross pilot.

The Barcelona pilot, in addition to validated questionnaires, will be in a position to capture data on patient QoL issues also through the Xemio app which patients will be encouraged to use to obtain information about any discomforts or other symptoms/side-effects they are experiencing offering the opportunity to collect data about QoL issues at the point when they are most pertinent to the patient in addition to the periodic self-reporting based on validated

questionnaires. The Xemio app data collection interface will be based on simple questions the patient will not find it inconvenient and tiresome to answer.

The CareAcross questionnaires are also designed to be shorter and simpler than the validated questionnaires used in other pilots. This reflects an understanding that patients voluntary supplying data to a remote online platform, the CareAcross website, are far more unlikely to do so if this is tiresome and time-consuming. CareAcross's data collection is tied with the supply of content, an approach which will also be taken during ASCAPE.

Within these two exceptions (Barcelona Xemio app, CareAcross website), the simple questions will be aligned to the validated questionnaires used by the other pilots.

3.1.2.2 Wearables

Three pilots (Athens, CareAcross, Örebro) will use wearables to collect active monitoring data. All three pilots will use the same wearable model that has been selected according to the procedure described Section 2.5. The characteristics of the data to be collected from wearables are described in Annex I.

3.1.2.3 Mobile App

One pilot (Barcelona) will use a mobile app (Xemio, developed by iSYS Foundation) as a patient-driven tool for reporting more than 50 health issues. The app has also the ability to capture step count data (per day) as well as frequency and duration of app use. As a result, the Xemio app will be used to capture both relevant QoL issues and active monitoring data in the Barcelona pilot. In addition, Xemio has a social calendar in which patients can look for social activities and patient education events and log the activities they participate in.

3.1.2.4 Web-based platform***

CareAcross pilot is based on a web-based platform for collecting patients' responses to general questions (patient- and tumour characteristics, treatment strategy), specific questions related to side effects and QoL issues, and validated questionnaires regarding QoL issues.

3.1.2.5 Prospective Datasets

The prospective datasets will be built upon the data collected through validated questionnaires (all four pilots), wearables (Athens, CareAcross, Örebro), mobile app (Barcelona), and web-based platform (CareAcross).

For all except one (CareAcross) pilot, the prospective datasets will also be enriched by relevant data from Medical Records that will be captured through electronic Case Report Forms for each eligible patient.

3.2 Mapping the data types and definitions of QoL issues within datasets

The identification of QoL issues to be predicted within ASCAPE for breast and prostate cancer was a multistep procedure that was based on current evidence and clinical expertise.

Specifically, a literature search, using the electronic database PubMed, was first performed with the following keywords: QoL issues, QoL aspects, health issues, health problems, side effects, and breast cancer or prostate cancer. No year or language restriction was applied. However, the searching strategy was limited only to systematic reviews or meta-analyses to capture the cumulative evidence on the topic. The searching strategy yielded nearly 1,300 potentially eligible articles for breast cancer and 650 for prostate cancer. The first study selection was conducted on the title and abstract and the second on the full text. The eligible studies after the two-step selection process were used to build a list of potential QoL issues for breast and prostate cancer, respectively.

To ensure that the process captured all the potential QoL issues, we compared the lists from the searching strategy with the lists from CareAcross retrospective datasets based on patients' answers on a web-based platform about their actual QoL issues. After this process, 51 QoL issues for breast cancer and 47 for prostate cancer were identified and included in the lists.

The lists of QoL issues on breast and prostate cancer, respectively were sent to the clinical experts within each clinical partner with instructions to prioritise the QoL issues considering the following aspects: how frequent each QoL issue is, the magnitude of impact on patients' daily life, and the potential positive impact if the QoL issue can be predicted before it occurs.

For breast cancer, the prioritisation process (clinical experts from Barcelona, CareAcross, Örebro) identified 15 QoL issues (Table 3) to be predicted through ASCAPE-platform.

Table 3 Quality of life issues for breast cancer for prediction within ASCAPE

QoL issue	N of retrospective databases with relevant information
Anxiety	2
Body changes	0
Body image	1
Cognitive impairment	1
Depression	3
Dry vagina	0
Emotional symptoms (loneliness)	1
Fatigue	3
Hot flushes	2
Insomnia	3
Joint pain	2
Local symptoms after surgery	1
Lymphedema	3
Neurotoxicity	3
Sexual dysfunction	2

For prostate cancer, the prioritisation process (clinical experts from Athens, CareAcross, Örebro) identified 12 QoL issues (Table 4) to be predicted through ASCAPE-platform.

Table 4 Quality of life issues for prostate cancer for prediction within ASCAPE

QoL issue	N of retrospective databases with relevant information
Anxiety	1
Bowel dysfunction	2
Cognitive impairment	0
Depression	1
Erectile dysfunction	2
Fatigue	1
Hot flushes	1
Incontinence	1
Low urinary tract symptoms	2
Loss of libido	0
Musculoskeletal pain	0
Weight changes	1

3.2.1 Mapping the QoL issues in breast cancer within retrospective datasets

The process of mapping the 15 QoL issues in breast cancer within the five available retrospective datasets revealed that there was at least one dataset with relevant information for 13 QoL issues whereas no retrospective data was found for two QoL issues (body changes, dry vagina).

The collection strategies for the QoL issues within retrospective datasets varied and included validated questionnaires, simple questions with Boolean expressions, and proxy (new prescription of specific medications). As a result, the definition of the presence or absence of each QoL issue was also differed across datasets. The mapping process within retrospective datasets was, therefore, performed on two levels: the collection strategies for each QoL issue and the definitions of the presence or absence of each QoL issue.

The definitions applied at each QoL issue depend on the collection strategy as follows:

- Validated questionnaires as collection strategy: the presence of QoL issue was defined based on current evidence and clinical expertise considering the minimal clinical important difference (MCID) approach [11, 12].
- Simple questions with Boolean expressions as collection strategy: the presence of QoL issue was defined using the Boolean expression (yes OR true for present / no OR false for absent)
- Proxy as collection strategy: the presence of QoL issue was defined as any new prescription of relevant medication (e.g. antidepressants for depression) for at least 2 consecutive prescriptions after breast cancer diagnosis.

For each QoL issue, the definitions are considered equal among different retrospective datasets, e.g. for insomnia, the patients will be considered as having problem with insomnia if they have a MCID according to the Fibromyalgia Impact Questionnaire (Sörmland retrospective dataset) OR true to the simple Boolean question (CareAcross retrospective dataset) OR 2 consecutive prescriptions of any medication against insomnia after breast cancer diagnosis (BcBaSe retrospective dataset).

A summary of mapping process within retrospective datasets is presented in Annex I, Table 17.

3.2.2 Mapping the QoL issues in breast cancer between retrospective and prospective datasets

Among the three prospective datasets (Barcelona, CareAcross, Örebro), data for seven QoL issues will be collected from all three datasets, three from two datasets, and seven from one dataset.

The collection strategies will include either validated questionnaires or simple

questions with Boolean expression.

The definitions for each QoL issue depend on the collection strategy with the same principles as described above:

- Validated questionnaires as collection strategy: the presence of QoL issue was defined based on current evidence and clinical expertise considering the minimal clinical important difference (MCID) approach [11, 12].
- Simple questions with Boolean expressions as collection strategy: the presence of QoL issue was defined using the Boolean expression (yes for present / no for absent).

For each QoL issue, the definitions are considered equal among different prospective datasets, e.g. for insomnia, the patients will be considered as having problem with insomnia if they have a MCID according to the EORTC QLQ30 questionnaire (Örebro and Barcelona prospective dataset) OR true to the simple Boolean question (CareAcross) OR at least 2 to the 4 point Likert-scale simple question (Barcelona pilot).

When two different validated questionnaires will be used to collect information about the same QoL issue between different prospective datasets, the MCID approach will be separately applied to the two questionnaires and a common definition will be adopted.

For each QoL issue, the definitions between the retrospective and prospective datasets are considered equal according to the above-mentioned principles.

A summary of mapping process among prospective datasets is presented in Annex I, Table 18.

3.2.3 Mapping the QoL issues in prostate cancer within retrospective datasets

The mapping process of the 12 QoL issues in prostate cancer within the two available retrospective datasets identified at least one dataset with relevant information for nine QoL issues whereas no retrospective data was found for three QoL issues (cognitive impairment, loss of libido, musculoskeletal pain).

The collection strategies for the QoL issues within retrospective datasets varied and included validated questionnaires and simple questions with Boolean expressions. As a result, the definition of the presence or absence of each QoL issue was also differed across datasets. The mapping process within retrospective datasets was, therefore, performed on two levels: the collection strategies for each QoL issue and the definitions of the presence or absence of each QoL issue.

The definitions applied at each QoL issue depend on the collection strategy as follows:

- Validated questionnaires as collection strategy: the presence of QoL issue was defined based on current evidence and clinical expertise considering the minimal clinical important difference (MCID) approach [11, 12].
- Simple questions with Boolean expressions as collection strategy: the presence of QoL issue was defined using the Boolean expression (yes OR true for present / no OR false for absent)

For each QoL issue, the definitions are considered equal among different retrospective datasets, e.g. for low urinary tract symptoms (LUTS), the patients will be considered as having problem with LUTS if they have a MCID according to the International Prostate Symptom Score questionnaire (Örebro prostate cancer retrospective dataset) OR true to the simple Boolean question (CareAcross retrospective dataset).

A summary of mapping process within retrospective datasets is presented in Annex I, Table 19.

3.2.4 Mapping the QoL issues in prostate cancer between retrospective and prospective datasets

Among the three prospective datasets (Athens, CareAcross, Örebro), eight QoL issues will be collected from all three datasets whereas four from two datasets. The Athens and Örebro prospective studies will use the same questionnaires, while the CareAcross pilot will use either validated questionnaires (other than the ones from Athens/Örebro) or simple questions with Boolean expression.

The collection strategies will, therefore, include either validated questionnaires or simple questions with Boolean expression.

The definitions for each QoL issue depend on the collection strategy with the same principles as described above:

- Validated questionnaires as collection strategy: the presence of QoL issue was defined based on current evidence and clinical expertise considering the minimal clinical important difference (MCID) approach [11, 12].
- Simple questions with Boolean expressions as collection strategy: the presence of QoL issue was defined using the Boolean expression (TRUE for present / FALSE for absent).

For each QoL issue, the definitions are considered equal among different prospective datasets, e.g. for LUTS, the patients will be considered as having problem with LUTS if they have a MCID according to the EORTC PR25 questionnaire (Athens and Örebro prospective dataset) OR true to the simple Boolean question (CareAcross).

When two different validated questionnaires will be used to collect information about the same QoL issue between different prospective datasets, the MCID approach will

be separately applied to the two questionnaires and a common definition will be adopted.

For each QoL issue, the definitions between the retrospective and prospective datasets are considered equal according to the above-mentioned principles.

A summary of mapping process among prospective datasets is presented in Annex I, Table 20.

3.3 Identifying indicators for QoL issues

Considering the lack of evidence on AI-based approaches in monitoring QoL in cancer patients, the choice of potential indicators for QoL issues within ASCAPE was intentionally wide including all medically relevant information to make sure that all potentially predictors will be available for AI-based modelling.

The medically relevant information to be captured as potential indicators for QoL issues can be categorised as follows: patients' socioeconomic status, baseline characteristics on patients' lifestyle factors, baseline characteristics on patients' health condition, baseline characteristics on patients' reproductive factors (for breast cancer), family history of cancer, tumour characteristics (clinical and pathological features), treatment strategies (including type of treatment, duration, dates), baseline status on QoL issues.

In addition, active monitoring data (through wearables or mobile app) and actual nutritional data (through food frequency questionnaires (FFQ) and/or anthropometric measurements such as BMI during follow-up) will also be captured and included as potential indicators for QoL issues.

No laboratory results will be considered as QoL indicators because the current guidelines recommend against routine laboratory testing during the follow-up period in breast and prostate cancer with the exception of PSA for prostate cancer recurrence [13, 14, 15].

3.3.1 Indicators for QoL issues in breast cancer

Considering all the medically relevant information that could potentially be of interest as indicators for QoL issues in breast cancer, 72 unique indicators were identified. In addition, active monitoring data collected from wearables (or mobile app for Barcelona pilot) and nutritional data will also be included as potential indicators.

3.3.1.1 Availability and mapping of indicators within retrospective datasets

The availability of the potential indicators within retrospective datasets varies from 21% to 61%.

A mapping process has been started and will be completed during the WP2 (Task 2.1) to align the indicators across the retrospective datasets.

3.3.1.2 Indicators beyond retrospective datasets

There are several potential variables of interests that have not been collected among the retrospective datasets. These indicators can be divided into three categories: medically relevant information not available in retrospective datasets; active monitoring data; nutritional data.

These indicators are planned to be collected during the prospective studies through different collection strategies including electronic Case Report Forms, questionnaires, wearables, mobile app, and web-based platform.

3.3.1.3 Mapping the indicators between retrospective and prospective datasets

An extensive mapping process has been started and will be completed during the WP2 (Task 2.1) to align the nature of indicators within retrospective datasets with the way the same indicators will be collected during the prospective studies. In total, 60 of 72 indicators have been found to be the same between at least one retrospective dataset and prospective datasets.

3.3.1.4 Indicators *collected* from wearables

Following the procedure described in Section 2.5, seven parameters will be collected from wearables as potential indicators for QoL issues including: steps, activity time, activity type, calories burned, heart rate, sleep quality, sleep quantity.

The Barcelona pilot will use a mobile app instead of wearables and will be able to capture data on steps.

3.3.1.5 Collection strategies in prospective studies

Five approaches for prospective data collection will be used across the pilots: medically relevant information from Medical Records obtained through electronic Case Report Forms for each eligible patient (Barcelona, Örebro); validated questionnaires (Barcelona, CareAcross, Örebro), wearables (CareAcross, Örebro), mobile app (Barcelona), and web-based platform (CareAcross).

The way of collecting prospective data depends on the pilot design but also on the nature of the indicators to be collected.

3.3.2 Indicators for QoL issues in prostate cancer

Considering all the medically relevant information that could potentially be of interest as indicators for QoL issues in prostate cancer, 76 unique indicators were identified.

In addition, active monitoring data collected from wearables and nutritional data will also be included as potential indicators.

3.3.2.1 Availability and mapping of indicators within retrospective datasets

The availability of the potential indicators ranged from 25% (CareAcross retrospective dataset) to 55% (Örebro retrospective dataset).

A mapping process has been started and will be completed during WP2 (Task 2.1) to align the indicators across the retrospective datasets.

3.3.2.2 Indicators beyond retrospective datasets

Several indicators, including medically relevant information, active monitoring data, and nutritional data, have not been collected through the retrospective datasets.

These indicators are planned to be collected during the prospective studies through different collection strategies including electronic Case Report Forms, questionnaires, wearables, and web-based platform.

3.3.2.3 Mapping the indicators between retrospective and prospective datasets

An extensive mapping process has been performed to align the nature of indicators within retrospective datasets with the way the same indicators will be collected during the prospective studies. In total, 42 of 76 indicators have been found to be the same between at least one retrospective dataset and prospective datasets.

3.3.2.4 Indicators collected from wearables

Following the procedure described in Section 2.5, it was decided that seven parameters will be collected from wearables as potential indicators for QoL issues including: steps, activity time, activity type, calories burned, heart rate, sleep quality, sleep quantity. The parameters selected for QoL issues in prostate cancer are the same as in breast cancer.

3.3.2.5 Collection strategies in prospective studies

Four approaches for data collection will be used across the pilots: medical-relevant information from Medical Records obtained through electronic Case Report Forms for each eligible patient (Athens, Örebro); validated questionnaires (Athens, CareAcross, Örebro), wearables (Athens, CareAcross, Örebro), and web-based platform (CareAcross).

The way of collecting prospective data depends on the pilot design but also on the nature of the indicators to be collected.

3.4 Interventions for personalized support

The identification process for suitable QoL issues for proposing patient-centric interventions through the ASCAPE-platform was similar to the process for identification of QoL issues for prediction described in Section 3.2 with some differences on the aspects that were considered during the selection process.

Specifically, the initial lists of QoL issues for breast and prostate cancer were derived from the same searching strategy and algorithm as described in Section 3.2. The lists were sent to the clinical experts within each clinical partner with instructions to prioritise the QoL issues regarding the value of proposing patient-centric interventions through ASCAPE considering the following aspects: how frequent each QoL issue is found in the population, the ability for self-reporting, the potential need of clinical examination for proper evaluation, the number of potential interventions for each QoL issue, the timeframe for intervention needed for each QoL issue (urgent vs. non-urgent), and the anticipated risk of using AI for each QoL issue.

Through the above-mentioned selection process, 7 QoL issues for breast cancer patients and 7 for prostate cancer were considered suitable for proposing interventions through the ASCAPE-platform (Table 5).

Table 5 QoL issues suitable for proposing interventions through ASCAPE

Type of cancer	QoL issues
Breast cancer	Anxiety Depression Joint pain Fatigue Neurotoxicity Hot flushes Weight gain
Prostate cancer	Anxiety Depression Fatigue Incontinence Weight changes Sexual dysfunction Hot flushes

After prioritisation and identification of the suitable QoL issues to intervene with through the ASCAPE, a second process took place to categorise the QoL issues regarding the way to present the interventions in the ASCAPE-platform. We are planning to use two different ways: (a) a list of interventions for a specific QoL issue without priority among the different interventions and (b) a list of interventions for a specific QoL issue with priority based on current evidence. The reasons for this

categorisation rely on the fact that there is no available data on specific interventions and their efficacy in the available retrospective datasets and for some QoL issues there is also a lack of reliable evidence on which intervention is more suitable.

For QoL issues with potential interventions lacking reliable evidence, a list of interventions without any prioritisation will be proposed by the ASCAPE-platform. On the other hand, for QoL issues with reliable evidence regarding the potential interventions, a list of interventions with prioritisation based on the current evidence (according to the Oxford Centre for Evidence-based Medicine – levels of evidence) [16] will be proposed. The initial prioritisation of interventions will be done by the clinical experts within the ASCAPE based on a systematic review of the literature and a critical appraisal of published international guidelines on the management of QoL issues after breast or prostate cancer diagnosis. The initial prioritisation will be enriched by continuous machine learning on prospective data on the efficacy of interventions applied to the patients for each specific QoL in order to improve the ability of ASCAPE-platform to propose to healthcare professionals a prioritised list of interventions with higher confidence level.

4 Pilot Evaluation Framework

4.1 AI-Centric Evaluation

4.1.1 Evaluation measures for classification-based predictive models

Let *class* (*label*) be a single value of a categorical variable, and let *unclassified instance* be an instance that does not have a value for that variable, i.e. an instance that does not have a class. *Classification* is then a machine learning problem of assigning a class to an unclassified instance. The instance's class is determined by using a knowledge inferred from a dataset consisting of already classified instances. When a categorical variable has only two possible values, i.e. when there are in total two possible classes, the classification problem is called *binary classification*. On the other hand, when there are *n* possible classes, n being greater than two, the problem is called *n-ary classification*.

In the ASCAPE framework both binary and n-ary classification problems are expected to emerge. For example, binary classifications might emerge in the prediction of quality of life variables, where a variable could represent one concrete side effect, while its values would suggest if a patient has the side effect or not. N-ary classifications might emerge also in intervention suggestions, where multiple interventions are possible, while only one should be suggested for a given patient.

There are many algorithms that solve the classification problem; such algorithms are called *classifiers*. The performance of a classifier is highly influenced by the underlying dataset. Therefore, for a given dataset, multiple classifiers should be evaluated, and the best performing one should be used for further classifications. For that purpose, there exists various measures for classifiers evaluation. In the further text some well-known will be presented.

For the case of binary classification, one class is usually seen as a *positive* class, while the other is seen as a *negative* class. *True positives* (TP) is then a number of instances that are correctly classified with positive class, *false positives* (FP) is a number of instances that are incorrectly classified with positive class, *true negatives* (TN) is a number of instances that are correctly classified with negative class, and finally, *false negatives* (FN) is a number of instances that are incorrectly classified with negative class. Based on these numbers, three measures are derived: recall, precision and accuracy. They are calculated as follows:

$$recall = \frac{TP}{TP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

As it can be seen, *recall* is a ratio between the number of correctly classified positive instances and the number of all positive instances. Therefore, it should be used to determine the extent of positive instances that the algorithm managed to detect. On the other hand, *precision* is a ratio between the number of correctly classified positive instances and the number of all positively classified instances. Hence, the precision should be used to determine if the error among positively classified instances is small enough – higher precision value indicates lower error. Finally, *accuracy* is simply ratio between the number of correctly classified instances and the number of all classified instances, and it therefore captures overall classifier's accuracy.

Both recall and precision are very often equally important. In order to evaluate an algorithm considering both of them, F-measure (*F_1 measure*) could be used. F-measure is harmonic mean of recall and precision, hence it is calculated as follows:

$$F = \frac{2}{recall^{-1} + precision^{-1}}$$

A proper evaluation of binary classifiers should imply examination of all the presented measures, since each of them puts a light on different aspects of an algorithm. A few examples will be presented that demonstrate how looking only at a single measure could be very misleading. The classification algorithm that classifies all the instances positively will have the highest possible recall value, even though such algorithm is obviously not a good classifier. In the unbalanced datasets, in which most of the instances have the same class, the main objective is usually to detect the instances of minor class. In this problem setup, high accuracy definitely does not mean that the objective is met. Moreover, if the major class is the positive class, high precision as well becomes meaningless.

The presented measures cannot be used as such for the case of n-ary classification problem. However, they can be adapted for it. Let us first introduce the notion of the *confusion matrix*. The confusion matrix is a square matrix whose rows represent real instances' classes, while its columns represent classes assigned to the instances by a classification algorithm. The confusion matrix's element $M_{X,Y}$ (an element in row X and column Y) is equal to the number of instances of class X classified with class Y. Consequently, good algorithm should maximize the numbers on the main diagonal of the confusion matrix. Moreover, if high numbers outside of the main diagonal exist, they represent the specific places where the classifier misbehaves.

Recall, precision and F-measure can be defined for each class X by using the values from the confusion matrix in the following way:

$$recall_X = \frac{M_{X,X}}{\sum_Y M_{X,Y}}$$
$$precision_X = \frac{M_{X,X}}{\sum_Y M_{Y,X}}$$

$$F_X = \frac{2}{{recall_X}^{-1} + {precision_X}^{-1}}$$

Contrary to that, accuracy is defined on the level of the whole algorithm, as follows:

$$accuracy = \frac{\sum_X M_{X,X}}{\sum_{X,Y} M_{X,Y}}$$

Additionally, all the three measures that are defined on the class-level (recall, precision and F-measure) can be micro- and macro-averaged, in order to depict an overall performance. Macro-average is simply a mean of values obtained for each separate class. Micro-average takes into account distribution of instances over different classes – value of a single class contributes to the overall result proportionally to the portion of the instances assigned to that class. Let v_X be a value of a single measure (recall, precision or F-measure) for the class X, i_X be the number of instances of class X, and i be the total number of instances. Micro-average (μ_{micro}) and macro-average (μ_{macro}) values are then calculated as follows:

$$\mu_{micro} = \sum_X \frac{i_X}{i} v_X$$

$$\mu_{macro} = \frac{\sum_X v_X}{i}$$

All the presented measures will be used throughout all the testing and evaluation phases of the ACAPE federated learning (FL) processes that enable a range of federated ML models. Initially, they will be used in the evaluation of the various classification algorithms, with the final aim being selection of the most suitable algorithms for processing the ASCAPE Pilot partners' datasets. Afterwards, they will be used to evaluate the extent in which doctors' final decisions are aligned with the suggestions made by the ASCAPE platform. In that case, doctors' decisions will be viewed as instances' real classes.

4.1.2 Evaluation measures for regression-based predictive models

In machine learning, regression is used when labels take values in a continuous range. Hence, instead of classifying a sample as belonging to a specific class, the model will output a numerical value. Regression models are typically used in finance and biomedical engineering.

Regression models could be used in ASCAPE platform for prediction of side effects severity. Namely, for certain side effects it might not be enough to predict only their presence, but also the form in which they appear. Moreover, regression models might also be used to predict time frames in which certain side effects are expected to appear. Finally, regression models could also be employed in case of multi-class

classification problems, especially if numerical, ordered values are assigned to the classes (e.g. QoL metric evaluated on a scale from 1 to 10).

There are several typical functions which measure the prediction error of a regression model, which are described briefly in the following.

Mean Squared Error (MSE), also called Quadratic loss or L2, measures the average of the squares of the errors. The output is equal to the average square difference between the predicted values and the real ones, defined by the following formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

where n is the number of samples, Y is the vector of observed values of the variable being predicted, and $\hat{Y}$ is the vector of the predicted values. MSE is the most commonly used regression loss function.

Root Mean Squared Error (RMSE) is calculated as the square root of mean squared error. RMSE is sensitive to outliers and can often exaggerate results. Like MSE, the output will be always a positive value. RMSE is defined by the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

Mean Absolute Error (MAE) is another common loss function used for regression models. MAE is defined as the sum of absolute differences between ground truth and predicted values:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right|$$

It measures the average magnitude of errors in a set of predicted values, without considering their directions. Compared to RMSE, MAE penalizes large errors less. The RMSE result will always be larger or equal to the MAE. If all of the errors have the same magnitude, then RMSE is equal to MAE. If not, RMSE is equal to $\mathrm{MAE} * \sqrt{n}$, where n is the number of samples.

Relative Squared Error (RSE) is a metric computed as ratio between the sum of squared errors and the sum of squared differences between the ground truth values and their average value. RSE basically evaluates the error with respect to a baseline predictor which always outputs the average value of the dataset. RSE is defined by

the following formula:

$$RSE = \frac{\sum_{j=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{j=1}^{n}(\hat{Y}_i - \bar{Y})^2}$$

where Y_i is the value predicted by the model, $\hat{Y}_i$ is the ground truth value, and $\bar{Y}$ is the average of the ground truth values.

Relative Absolute Error (RAE) is computed similarly to RSE, but instead of squaring the differences in the denominator and numerator, their absolute value is taken:

$$RSE = \frac{\sum_{j=1}^{n}|Y_i - \hat{Y}_i|}{\sum_{j=1}^{n}|\hat{Y}_i - \bar{Y}|}$$

Standard deviation (SD) is a common measurement used in statistics, for analysing the dispersion or variation in a set of values. A small standard deviation indicates that the values from the analysed dataset are close to the mean, while a high standard deviation indicates that the values are spread out over a wider range. Standard deviation is defined as follows:

$$\boldsymbol{\sigma} = \sqrt{\frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}}(\boldsymbol{x_i} - \bar{\boldsymbol{x}})^2}{\boldsymbol{n-1}}}$$

where $\boldsymbol{\sigma}$ is the standard deviation, $\boldsymbol{x_i}$ represents the values from a dataset, $\bar{\boldsymbol{x}}$ is the mean value of the dataset and *n* is the total number of data points. The standard deviation is also equal to the square root of variance, but, unlike variance, SD is expressed in the same units as the dataset.

Pearson correlation coefficient (PCC) is a statistic measurement for assessing the linear correlation between two variables. The coefficient takes values between +1, which indicates a total positive linear correlation, and -1, which indicates a total negative linear correlation. A value close to 0 indicates the lack of a correlation. PCC is calculated as the covariance of two variables divided by the product of their standard deviation:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\boldsymbol{\sigma_X} * \boldsymbol{\sigma_Y}}$$

where $\rho_{X,Y}$ is the person correlation coefficient between the *X* and *Y* variables, *cov* is the covariance, $\boldsymbol{\sigma_X}$ is the standard deviation of *X* and $\boldsymbol{\sigma_y}$ is the standard deviation of *Y*. PCC is very sensitive to outliers. One single unusual observation will have a huge impact on the final value of the correlation. Hence, it is highly recommended to visualize the data in a scatter plot and remove the outliers, before computing the

Pearson correlation score.

4.1.3 Feature evaluation measures

Feature (i.e. variable, attribute) selection and feature extraction are the two most common feature evaluation techniques [17, 18]. The purpose of these techniques is to reduce dimensionality of input data in the sense of reducing the number of features important for classification/regression tasks. This reduction can have significant positive impacts on trained machine learning models including [19]:

- reduction of effects of course of dimensionality,
- shorter training/inference times of ML models,
- improved generalization of the model by reducing overfitting,
- simplification of ML models
- in some circumstances even improvement in ML models accuracy

Feature selection is the process of selecting a subset of relevant features from an original set of features. This process relies on the fact that the datasets used for ML usually contains features that are redundant or irrelevant for the desired classification/prediction task.

Generally, feature selection algorithm searches for the combination of features which maximizes some form of feature evaluation measure. In majority of cases in ML this measure could be any evaluation measure for classification-based predictive models (precision, recall, accuracy, F-measure, etc.) The simplest algorithm for feature selection could be to test any combination of features and compare the values of evaluation measures. However, this approach is computationally unfeasible due to an enormous volume of search space. To cope with this problem, three main groups of feature selection algorithms can be identified considering evaluation measure:

- **Wrapper methods.** These methods use the ML classification model to score feature subsets. For each combination of features, the model is trained and evaluated. Since the whole space of feature subsets cannot be covered, these algorithms usually use metaheuristics and different ways of pruning of the search space. The main advantage of these methods is that they commonly find the optimal feature set for a particular ML model. The disadvantages are that these methods are computationally demanding, and that they can get stuck in local optimums.
- **Filter methods.** These methods do not use desired ML model to assess features. Instead, they use some more general measure which is easy to compute but still useful for scoring the feature subset. Some of the common measures are: Pearson correlation coefficient, information gain, symmetrical

uncertainty, Relief algorithms etc. These methods usually give the full list of features sorted descending by chosen measure, and the number of selected features is selected afterwards. Filters are computationally more efficient than wrappers, but they do not produce the optimal subset of features for chosen ML model. Furthermore, these measures tend to select redundant features since they usually don't consider the relationships between features.

- **Embedded methods.** These methods try to combine the good sides of previous methods. They perform feature selection as part of the model construction process. Two most popular methods in this group are LASSO (Least Absolute Shrinkage and Selection Operator) and Ridge Regression. As expected, the computational efficiency of embedded methods is between filters and wrappers.

On the other hand, feature extraction methods try to build a new set of features on the basis of original features. The new set of features is gained by applying particular functions on the original set of features. Feature extraction presents one form of dimensionality reduction technique, and some of the common techniques are: Principal component analysis, Independent component analysis, Isomap, Autoencoder etc. Newly produced set of features tend to be non-redundant, but their semantic and interpretability is commonly questionable.

Feature selection techniques will be used during ASCAPE project in order to select features relevant for desired classification/regression tasks. These include both prediction of QoL issues and the proposition of health interventions needed for handling of predicted issues. Furthermore, feature selection will be applied on both retrospective data (data already collected by clinical partners) and on prospective data (data which will be collected during ASCAPE project). Feature extraction methods, especially Autoencoders, will be used for checking whether newly collected data significantly differs from the past data for the purpose of outlier detection.

4.1.4 Evaluation metrics for unsupervised models

Clustering or cluster analysis can be considered as the task of grouping a set of objects/instances (in any format), such that objects form the same group, i.e. cluster, are more similar to each other than those in the other clusters. Cluster analysis does not refer to a specific algorithm, but to the general task that to be solved and represents an unsupervised machine learning technique.

Cluster analysis will be used within ASCAPE project mainly for outlier detection. A common scenario might be as follows: the global model of ASCAPE framework should be updated with the new knowledge from new data. However, if it turns out that the prediction accuracy of new data is considerably low, then the new data will

be clustered. Subsequently, the accuracy of the model will be tested on each cluster. The clusters with low classification accuracies will be treated as outliers, and will be selected for further investigation, while only the clusters with high accuracies will be used for ASCAPE model update. In such a way, the AI model will be protected from intentional or unintentional data injection attacks with non-valid data.

Additionally, cluster methods might be used for various versions of data analysis within ASCAPE project. For example, clustering can be used for identifying different cohort of patients (patients in different stage of illness or treatment).

Silhouette index is a method used for interpretation and validation of consistency within clusters of data. In unsupervised learning, a challenging aspect is the choice of the number of clusters. For one dimensional and two-dimensional data visual representation can guide the decision, but such an approach becomes more difficult to follow when the number of dimensions is higher. There are two popular methods to tackle this challenge: Inertia and the Silhouette index. The Silhouette index is regarded as the better approach compared to Inertia. It is computed for each instance with the following formula:

$$Silhouette\ Coefficient\ = \frac{(x-y)}{max(x,y)}$$

where y is the mean distance to the other instances in the same cluster, and *x* represents the mean distance to the nearest cluster. This coefficient takes values between -1 and 1. A value close to 1 means that the instance is correctly associated to its cluster. In contrast, a value close to -1 means that the instance is not associated to the right cluster. This is a computation intensive algorithm since the coefficient is calculated for every instance.

The RAND index (or RAND measure) is a measure used in statistics and data clustering, representing the similarity between two clusters of data. It is defined by the following formula:

$$R\ = \frac{a+b}{n(n-1)/2} = \frac{a+b}{\binom{n}{2}}$$

where *a* is the number of times a pair of elements belongs to the same cluster, *b* is the number of times when a pair of elements are in different clusters, *n* is the total number of elements in a set and $\binom{n}{2}$ is the number of unordered pairs in that set.

The RAND index has values ranging from 0 to 1, where 0 indicates that the two data clusters do not agree on any pair of points, and 1 indicates that the data clusters are identical.

The RAND index can also be calculated as a measure of the percentage of correct

decisions made by the algorithm, with the following formula:

$$R = \frac{TP + TN}{TP + FP + FN + TN}$$

where *TP* is the number of true positives, *TN* is the number of true negatives, *FP* is the number of false positives and *FN* is the number of false negatives.

Mutual information (*MI*) of two random variables can be viewed as a measure of the mutual dependence between those two variables.

Normalized Mutual Information (*NMI*) is a normalization of the Mutual Information (MI) where the score is scaled between 0, which means no mutual information, and 1, which means perfect correlation. *NMI* is defined by the following formula:

$$NMI(Y, C) = \frac{2 * I(Y; C)}{[H(Y) + H(C)]}$$

where *Y* are the class labels, *C* the cluster labels, *H(.)* the entropy, and the mutual information. *NMI* can be also used to compare two clusters with a different number of clusters.

Another metric for comparing detected clusters with ground-truth clusters is called Adjusted Mutual Information (AMI), which is an adjustment of the Mutual Information (MI) to account for the chance. The disadvantage of MI is that it is directly proportional to the number of clusters, even if more information is shared. AMI is independent from the absolute values of the labels, such that permutations of the class of cluster labels won't change the final score value.

4.1.5 Model evaluation methodologies

A machine learning model should exhibit a high degree of generalization, which is the ability of the model to give correct predictions on new, previously unseen inputs (data instances or data points) that were not present in the dataset used to learn the model (the training dataset). In an ideal case, the model (or multiple competitive models for a task at hand) can be evaluated on an independent test dataset containing data instances expected in practice that were collected independently from the training dataset. In this way, we can measure two kinds of ML model errors (resp., accuracies) for a given ML model performance measure *P* reflecting the badness (resp., goodness) of the model (this can be any measure described in the previous subsections):

1. The training error (resp., accuracy): *P* computed on the instances from the training dataset.
2. The test (generalization) error (resp., accuracy): *P* computed on the instances from the test dataset.

Good ML models are those having both low training and test errors (resp., high training and test accuracies). On the other hand, poorly performing ML models are associated to two undesirable phenomena frequently occurring in machine learning practice known as underfitting and overfitting. Underfitting is characterized by a high training error (resp., a low training accuracy), while overfitting happens when the difference between the training and test error (resp., accuracy) is too large.

Slight modifications made to machine learning algorithms aiming to reduce the test error (resp., increase the test accuracy) while keeping the same level of the training error (resp., accuracy) are collectively known under the term regularization. Different regularization mechanisms (e.g., adding penalty terms to a cost function that is optimized by a ML algorithm, early stopping and dropouts in neural networks) can be applied to prevent overfitting and make ML models more robust to unknown inputs. On the other hand, the application of a more complex machine learning algorithm (complex in terms of a more broader hypothesis space, which is the space of functions that can be learned by the algorithm) is a common strategy to deal with underfitting issues.

If the machine learning algorithm forming the model has parameters (also called hyper-parameters) then they should be tuned to obtain a ML model satisfying the previously mentioned property of good ML models. An independent validation dataset (independent from both training and test datasets) is required to select the best configuration of hyper-parameters for a ML algorithm in an unbiased fashion. Thus, in the ideal case the training of competitive ML models of different types accompanied with their validation is performed using three independent datasets:

1. The training dataset that is used to train multiple ML models for different configurations of hyper-parameters for each model type.
2. The validation dataset that is used to identify the best configuration of hyper-parameters for a particular model type and select the best model from the training phase per model type.
3. The test dataset: the dataset used to compute the generalization error (accuracy) for different model types and select the best ML model.

In practice, however, only one dataset is generally available to ML practitioners. The simplest strategy to deal with such situations, known as holdout, is to make a random division of the dataset into fixed training, validation, and test datasets. Usually, 20% or 33% of data instances in the dataset are randomly selected for the test dataset. The rest of the dataset is then divided into the training and validation datasets, where 20% of the instances are used in the validation phase and 80% in the training phase. Holdout random divisions should be done in a stratified way, which means that values of the outcome variable (the variable predicted by the ML model) should be equally represented in the training, validation and test datasets.

The holdout strategy can be problematic if the dataset is too small. A small test set implies a large statistical uncertainty regarding the generalization error (resp., accuracy). Consequently, more statistically robust approaches for ML model

evaluation were devised in the literature, including cross-validation and bootstrapping as the two most used approaches.

The main idea of bootstrapping is to perform a random selection with replacement of data instances from a given dataset to form the training dataset (and split it after into the training and validation datasets). The random selection with replacement means that one instance can be selected multiple times, appearing repeatedly in the training dataset. Let *N* denote the number of instances in the given dataset. If *N* random selections with replacement are performed over the dataset then approximately 63,2% of instances are going to be present in the training dataset (which will contain *N* instances, some of them will be repeated). The instances not selected by the previously described sampling procedure will be assigned to the test dataset. The generalization error (resp., accuracy) is then computed as

$$P = 0.632 \cdot P(\text{test}) + 0.368 \cdot P(\text{training}),$$

where *P*(test) is the value of *P* computed on the test set and *P*(training) is the value of *P* computed on the training dataset (also including the instances from the validation dataset). The bootstrapping procedure can be repeated several times and the final estimate of *P* is then computed by averaging obtained *P* values.

The cross-validation procedure involves splitting the given dataset into (stratified) non-overlapping folds. The number of folds is fixed, so this technique is also known as the *k*-fold cross-validation, where *k* is the number of folds (usually 5 or 10). The *k*-fold cross-validation procedure is done in *k* steps, each of them involving training and test phases for a fixed configuration of hyper-parameters, resulting with *k* estimates of *P*, which are then averaged into the final estimate of *P*. In the *i*-th step, *i*-th fold is used for testing, while the rest of folds are used for training. The whole procedure is described in the pseudocode given in Figure 3.

```
K-fold cross-validation(D, P, M):
inputs:
- D: dataset
- P: ML evaluation measure
- M: ML training algorithm (with fixed hyper-parameters)
output: estimate of P

randomly split D into k stratified folds D1, D2 to Dk
PV = an empty array of real numbers of size k

for i = 1 to k do
    T = Di
    R = D \ Di
    mod = train a model on R using M as the learning algorithm
    PV[i] = evaluate mod by computing P of mod on T
end

return average(PV)
```

Figure 3 K-fold cross-validation described in pseudocode.

The *k*-fold-cross validation procedure can be executed several times for different random splits into *k* folds. Leave-one-out cross-validation is the frequently used special case of k-fold cross-validation for k = N, where N is the number of instances in the dataset. This means that in each iteration of the leave-one-out cross-validation the test set contains exactly one data instance. The nested k-fold cross-validation is

```
Nested k-fold cross-validation(D, P, M, C):
inputs:
- D: dataset
- P: ML evaluation measure
- M: ML training algorithm
- C: set of hyper-parameter configurations for M
output: estimate of P, the best configuration of hyper-parameters from C

randomly split D into k stratified folds D1, D2 to Dk
PV = an empty array of real numbers of size k
LC = an empty array of hyper-parameter configurations of size k

for i = 1 to k do
    T = Di
    R = D \ Di
    randomly split R into k stratified folds F1, F2 to Fk

    tbc = None          /* the best configuration of hyper-parameters */
    tbcp = None                                   /* the P value of tbc */
    foreach c in C do
        PC = an empty array of real numbers of size k
        for j = 1 to k do
            V = Fj                                    /* validation dataset
*/
            W = R \ Fj                                  /* training dataset
*/
            mod = train a model on W using M(c) as the learning
algorithm
            PC[j] = evaluate mod by computing P of mod on V
        end
        pc = average(PC)
        if pc better than tbcp then
            tbc = c
            tbcp = pc
        end
     end

     mod = train a model on R using M(tbc) as the learning algorithm
     PV[i] = evaluate mod by computing P of mod on T
     LC[i] = tbc
end

return average(PV), mode(LC)
```

an extension of the k-fold cross-validation with an inner k-fold cross-validation loop enabling the construction of validation datasets and selection of best hyper-parameters. This procedure is described in the pseudocode given in Figure 4.
Figure 4 Nested k-fold cross-validation described in pseudocode.

4.2 Doctor-Centric Evaluation

The evaluation of AI-based innovations in healthcare regarding physicians' views and experience is challenging and needs to consider different aspects. The doctor-centric evaluation of ASCAPE will be focused on three axes:

1. Impact of ASCAPE on relevant metrics in clinical practice.

   Within this aspect, the following evaluation metrics (gathered as physicians' views and experience by using ASCAPE) will be considered: improvement in patient-doctor relationship; ASCAPE's efficiency to capture relevant QoL issues on time; changes in management or referrals made due to ASCAPE; usefulness of the information provided by ASCAPE; acceptability of integrating ASCAPE services into clinical practice; assessment of the time needed to use ASCAPE in clinical practice

2. Interaction between ASCAPE and physicians

   This aspect includes issues related to the interaction between the ASCAPE platform and physicians as usability, accessibility, and qualitative assessment of the interface.

3. Experience using the ASCAPE platform

   This aspect includes the more general issues on physicians' experiences in using the ASCAPE platform as trustworthiness, how confident physicians are regarding the reliability of ASCAPE, and psychological aspects in using an AI-based platform in clinical practice as perceived substitution crisis and behavioural intention.

Considering the complexity of the doctor-centric evaluation process, both quantitative and qualitative approaches will be used to ensure a wide coverage of evaluation metrics.

A detailed description of doctor-centric KPIs, derived from the above-mentioned evaluation axes, and the expected results are presented in Table 6.

Table 6 KPIs for doctor-centric evaluation of the ASCAPE

Doctor-centric evaluation axis	KPI title	KPI description	Expected result
ASCAPE impact in clinical	Doctor perception of their	The doctor will be asked to provide an	>70% of doctors will report an improvement in doctor-patient relationship

practice	relationships with the patient before and after use of ASCAPE platform	answer regarding the improvement or decline in level of their relationships with patients	
ASCAPE impact in clinical practice	Efficiency to capture relevant QoL issues in time (ability to capture a QoL issue earlier than the conventional follow-up)	The doctor will be asked to provide an estimated difference in detecting a QoL issues when using ASCAPE compared to conventional follow-up	At least 50% of the doctors reporting an earlier detection of QoL issues for at least one QoL issue is anticipated
ASCAPE impact in clinical practice	Evaluation of the interventions proposed by ASCAPE	The doctor will be asked whether the interventions proposed by ASCAPE are relevant and in accordance with clinical practice	>80% of the doctors will agree with the relevance of the proposed interventions
ASCAPE impact in clinical practice	Doctor perception of the usefulness of information provided by ASCAPE	The doctor will be asked if the predictions and interventions provided by ASCAPE is useful	>80% of the doctors will classify the information as moderate to very useful
ASCAPE impact in clinical practice	Doctor experience in integrating ASCAPE platform during everyday clinical practice	The doctors will be asked whether they will accept the ASCAPE platform into clinical practice	>80% of the doctors will be positive regarding the acceptance of ASCAPE into the clinical practice
Interaction between ASCAPE and physicians	Doctor perception on the accessibility (referred to ease of use and apply of ASCAPE-	The doctors will be asked to provide information on the ease of use and apply of the ASCAPE-platform	>75% of doctors will classify the platform as easy to use

	platform)		
Interaction between ASCAPE and physicians	Usage of ASCAPE by the Doctor	The ASCAPE system will be able to provide data on the number of times that they have used ASCAPE/total number of times they could have used it with their patients (i.e. the doctor used it on 80% of patient visits that the doctor has in the ASCAPE dashboard)	Doctor uses ASCAPE platform in >80% of the visits.
Interaction between ASCAPE and physicians	Doctor perception of ASCAPE interface quality	The doctor will be asked to provide an assessment of interface quality.	>85% of doctors will assess the ASCAPE interface as moderate to high quality
Experience using ASCAPE	Doctor experience regarding ASCAPE platform trustworthiness	The doctor will be asked to provide their experience on trustworthiness to use the platform for predictions and interventions	>80% of doctors will experience a moderate to high level of trustworthiness for both predictions and interventions
Experience using ASCAPE	Doctors' confidence in using ASCAPE	The doctors will be asked to provide their level of confidence in using ASCAPE for predictions and intervention	Moderate to high level of confidence for at least 80% of the doctors
Experience using ASCAPE	Doctor perception when shifting clinical practice from traditional to AI approach	The doctors will be asked to provide how they feel regarding the shift of clinical	>70% of the doctors will be positive on the shift of clinical practice; >70% of the doctors will be positive to use ASCAPE in the future

	and behavioural intention	practice from the traditional to an AI- based approach while using ASCAPE platform and their intention in using such approach on the future	

4.3 Patient-Centric Evaluation

In modern medicine, the focus on the aspects to measure improvement in the quality of healthcare processes has been based on different approaches during the last 50 years. First, Donabedian defined quality in Medicine with an emphasis on structure, process, and results [20]. Later, Cochrane emphasized effectiveness [21] and lastly, the Institute of Medicine focused on safety [22]. All these theories are valid paradigms regarding quality of healthcare but have all been defined from a scientific perspective. Currently, we are finding ourselves more interested in evaluating patients' perspective as an important aspect on quality in Medicine. Muir Gray [23] identifies a paradigm shift in the definition of quality towards value, understood as the standardization of the measurement of health improvements. Health care is satisfactory if it adds value from the patient's perspective, so the definition must always be done together with the patient.

The current approach is, thus, to add the patients' perspective to the already existing scientific perspective of evaluating quality in healthcare processes. To quantify the patients' perspective in order to evaluate healthcare processes Patient Reported Experience Measurements (PREMs) are used [24, 25].

Apart from the PREMs, the PROMs [26] (Patient Reported Outcome Measurements), referring to the measurement of patients' perception of their health, is another aspect of interest in supporting patient-centric care. Interestingly, using digital platforms to collect PREMs and PROMs have been shown to be highly effective.

Evaluating innovations, as ASCAPE, is quite challenging, and this is even more challenging in the case of healthcare. Moreover, ASCAPE innovations are related to the quality of life of patients with cancer, and this complicates the matter even further.

Cancer patients' experiences with healthcare systems overall consists of a variety of touchpoints, not all of which carry the same weight regarding their medical

outcomes. At the same time, their evaluation of their “cancer journey” as it relates to these systems is a very personal and often subjective one, because it may be affected by individuals (healthcare professionals or others) as well as operational aspects (from parking availability to hospital food etc).

While we acknowledge the idiosyncrasies of the patient experience and their evaluation, ASCAPE aims to evaluate patients’ experience from using the benefits and innovations of ASCAPE, in multiple ways using PROMs, PREMs, and measures of effectiveness.

Engagement and PROMs

During the pilot studies, patients' adherence / engagement to fill out questionnaires and provide data through wearables or Apps is essential as an indirect evaluation of the level of ASCAPE use.

Regarding questionnaires, the engagement will be measured either through the analysis of direct entries to the ASCAPE platform (for cases with direct entry) or through the analysis of indirect entries to the ASCAPE platform (for cases with clinician interaction)

Regarding engagement with wearables and App, there are recognized and applied theories that quantify the engagement of the use of technology such as the Technology Acceptance Model (TAM) [27]. TAM affirms that a technology must be easy and useful in order to be accepted. In addition, the Theory of Planned Behaviour postulates that the likelihood that an individual will engage in healthy lifestyle (e.g., regular exercise) is related to his/her motivation [28]. Lastly, the Fogg behaviour model shows that three elements must converge at the same time for a given behaviour to occur: motivation, ability (knowing how to do it), and action. When the desired behaviour does not occur, at least one of these three elements is missing [29]. Based on these models, the wearables and apps must be easy to use, useful, and the individual should show interest in using them in order to achieve a regular use. An indirect measure of TAM is the time these systems are used. Therefore, the evaluation of the acceptance of wearables and apps from the patients will be done through capturing the number of questionnaires, side effects, and QoL issues reported to the App.

Experience and PREMs

Patients’ experience to ASCAPE-based services is an important aspect in evaluation the ASCAPE. To quantify patients’ experience to ASCAPE, PREM-questionnaires will be used.

Measures of effectiveness and QoL improvements

The effectiveness of ASCAPE-based services is the ultimate measure of patient-centric evaluation of the ASCAPE. The effectiveness will be measured through the analysis of QoL issues where an improvement of QoL aspects is anticipating as a measure of the effectiveness of ASCAPE.

An additional measure of ASCAPE effectiveness will be the percentage of ASCAPE-suggested interventions patients report as following.

The patient-centric KPIs, derived from the above-mentioned evaluation aspects, are presented in Table 5.

Table 5. KPIs for patient-centric evaluation of ASCAPE

Patient-centric evaluation axis	**KPI title**	**KPI description**	**Expected result**
Engagement and PROMs	Questionnaires submitted per patient	The number of questionnaires submitted per patient during the follow-up	At least 75% of the planned questionnaires
Experience and PREMs	PREM-questionnaire	Patients will answer a summative PREM-questionnaire to capture the experience in using ASCAPE services	> 75% of the respondents will be classified as moderately to highly satisfied
Measures of effectiveness	Adherence to ASCAPE interventions	An analysis on whether the patients are following the intervention recommended by ASCAPE (in the cases where the doctor agrees with it and recommends it to the patient) Note: The extent to which doctors find ASCAPE intervention recommendations appropriate is captured by the	>70% of interventions suggested to doctors by ASCAPE and recommended to patients by their doctors will be followed

		corresponding Doctor-Centric Evaluation KPI	
Measures of effectiveness	Assessment of QoL	The QoL will be assessed through questionnaires during the use of ASCAPE services and an improvement in QoL status is anticipating	At least 20% QoL improvement at the end of time period where ASCAPE services will be available

4.4 Health Economics-Centric Evaluation

Around the world, healthcare providers are battling ever increasing healthcare demands and cuts to their budgets. In western developed countries that development is speeding up even more due to aging societies and associated conditions, including cancer. Altogether, this implies that healthcare managers must consider thoroughly any investment in healthcare innovation. Health economics assessment are fundamental in order to decide upon investments.

Such an evaluation is considered essential also for ASCAPE, as it will provide evidence of the benefits of ASCAPE to healthcare managers and facilitate the project exploitation efforts. The aim is to assess the health economics benefits of the project's services, considering a multitude of factors and relying on a methodology for a systematic evaluation of the services as well as the impacts of ASCAPE in terms of costs and healthcare outcomes. The relevant work will be based on information from the pilots and aim to estimate, at least theoretically, the value of the ASCAPE innovation in the countries of the four pilot sites: Greece, Spain, Sweden and UK. Altogether, it will provide valuable information not only for technical and clinical refinement but also for providing suitable business models based on cost-effectiveness evidence.

The methodology to be used will be described in D1.4 and the analysis of the data will be performed by using MAFEIP, a tool created to support evidence-based decision-making processes in the health and care sector [30]. MAFEIP is a web-based tool based on Markov model approach for economic evaluation and complies with the principles of Decision Analytic Modelling (DAM). We will use MAFEIP web-based tool to assess the impact of ASCAPE innovations in terms of health outcomes and economic resource used.

It is necessary to highlight that the proposed health economics study of the ASCAPE platform would ideally be performed by designing the pilots with two groups: a control group using standard care and an intervention group benefiting from ASCAPE. The ASCAPE breast and prostate cancer studies are not randomised trials with a control and an intervention group as this would be more appropriate for a second round of assessment of the efficacy of ASCAPE in a follow-up project or independent follow-up studies. However, all three clinical partner-lead pilots aim not only to evaluate ASCAPE but also to collect data before the First ASCAPE Prototype becomes operational in order to assist in the evolution of the ASCAPE models, their initial internal evaluation and their optimisation. D1.4 and D4.2 will address how the availability of data points with and without the use of ASCAPE can be used in order to conduct the health-economics evaluation in the clinic-lead pilots, as well as how it is to be conducted in the case of the CareAcross pilot.

Full details on the MAFEIP tool and the health economics-centric evaluation methodology will be provided in a dedicated deliverable: D1.4 (M10; outcome of Task 1.5). The health economics evaluation of ASCAPE will be carried out in WP5 and will be incorporated in D5.4; it will be conducted on the basis of the evaluation framework described in D1.4, the collection of relevant data about the performance of ASCAPE during the pilots, as well as a series of additional data (e.g. wearables cost, implementation costs, human resources costs etc.).

.

5 Conclusion

Demonstrating the potential of ASCAPE in a realistic setting requires the excellent collaboration of the technical and the pilot partners. The present deliverable demonstrates the readiness and ambition of both, but focuses on the contribution of the latter.

The ASCAPE pilots, as described in the pilot's design (Section 2), have undertaken to provide the testing ground for the evaluation of ASCAPE technologies in a variety of settings:

- in a clinical setting in the Athens and Orebro pilots
- in a setting which spreads beyond the hospital, to a regional network of primary healthcare centres in the Barcelona pilot, and
- in a European network of remotely supported patients in the CareAcross pilot.

A comprehensive evaluation framework for ASCAPE within the context of the four pilots was provided in Section 4. It covers AI-centric evaluation of the algorithms to be developed in WP2, the evaluation of ASCAPE from the doctors' and patients' point of view as planned for WP4, and the evaluation of ASCAPE from a health economics perspective as will be further developed in D1.4 and WP5.

On the technical side, the pilots also utilise different data collection and patient interaction modalities using technological means including a specialised mobile application (Barcelona pilot), a wearable device (Athens, CareAcross and Orebro pilots), web sites (Athens, CareAcross), email (Barcelona, Orebro). They will also demonstrate different levels of integration with ASCAPE: two pilots will only modify their IT systems in order to feed ASCAPE with patient data (CareAcross, Orebro), thus having to rely on the ASCAPE Dashboard to offer ASCAPE Functionality to doctors, whereas the remaining two (Athens, Barcelona) will attempt to provide full integration, of the kind envisaged in the exploitation scenario of ASCAPE. These differences were anticipated in the ASCAPE framework requirements (Deliverable 1.1) and will be catered for in the ASCAPE architecture (Deliverable 1.3), which, in turn, will serve as the starting point of WP3.

The principal challenge in the present deliverable addresses is the lack of sufficiently rich data sets for the purposes of the pilots being available at the outset. Given their focus on QoL, the pilots will need sufficient relevant data points from breast and prostate cancer patients for the relevant ASCAPE AI models to be trained. Ideally, large quantities of relevant data would have been available before the pilots began and efforts would concentrate on optimising what ASCAPE can do and how well using the available data. This is not the case. Despite the fact that numerous studies on QoL of cancer patients have been conducted, their data are not, as a rule, shared and additionally the data are not homogeneous or large-scale either. It is this reality that encourages initiatives such as the Open Research Data Pilot. However, in the

medical domain, there are significant challenges in sharing non-aggregated anonymised patient data, let alone making them public. As a result, any knowledge obtained from data collection in a study is restricted to one or more scientific publications focusing on general trends and not made available to AI algorithms that could apply it to specific patient cases. ASCAPE's innovative disruption of the status quo is that it allows knowledge from data collected at one site, either as part of a study or as part of a healthcare provider's standard protocols of patient monitoring, to be captured in AI models that can be shared and applied to individual patient cases. But as things stand, the ASCAPE pilots will have to grapple with the constraints of limited data sharing and the lack of a prior AI knowledge sharing platform which ASCAPE and the pilots would be able to build on and improve. ASCAPE pilots will engage in data bootstrapping initiated with retrospective datasets collected from the pilot partners or via their intervention from wider research networks in which they belong coupled with prospective data collected during the pilots used to continuously update the AI models. As the data for ASCAPE knowledge increases and predictions and intervention suggestions can be made with greater confidence, ASCAPE functionality available to the doctors will gradually increase also during the ASCAPE pilots. Importantly, ASCAPE functionality will remain available for healthcare providers that adopt ASCAPE as well as any follow up projects, continuously growing and providing an ever more solid foundation for personalised predictions and intervention suggestions.

The vision of creating AI knowledge from ever richer datasets coming from different healthcare providers requires an ability to use of heterogenous data. This challenge is not a theoretical challenge for future exploitation plans for ASCAPE; the ASCAPE pilots will have to be able to build the initial ASCAPE AI knowledge from a diverse set of retrospective datasets containing different data determinants and using different scales for the same data determinants. Those initial, retrospective, datasets will also be significantly different to the prospective datasets collected during the pilots, as the latter will not only include QoL indicators but also:

- data about interventions suggested by doctors,
- whether the intervention suggestions were followed and
- what their effect was on the patient's QoL.

The breast and prostate cancer prospective datasets, despite elements of common design such as the above, display various aspects of heterogeneity similar to those found in retrospective data. ASCAPE embraces differences in available information from different healthcare providers, as it is not only an inevitable factor in any future exploitation scenario, but also within the course of the ASCAPE pilots. The results of the pilots will depend to a large extent on how successfully this challenge is met.

In conclusion, the pilots do not present an artificially favourable testing ground for ASCAPE. The execution of the pilots in tandem with the efforts to complement the retrospective data with prospective data, and optimising the ASCAPE AI results that can be obtained from their combination represents a conglomeration of challenges that will both test the ASCAPE Consortium's capabilities and provide valuable

experiences for future follow-up research and exploitation.

Annex I. Wearables Requirements

Final Requirements List

Table 7 Wearables Requirements by Pilot Partners

Data-specific	Non data-specific
1. Security, Privacy and Anonymity, on the device 2. Security, Privacy and Anonymity, on the cloud or repository which is retrieving data collected on the device 3. Collection of data related to the Patient Needs for which Predictions are made (see Table 8) 4. Data collection that is as frequent as required by the algorithms to perform their Predictions 5. Data collection should not stop if communications interface (e.g. Bluetooth) is unavailable 6. Memory for data collected and not synchronized (e.g. due to communications interface being unavailable) should last at least 48 hours 7. Data collection that leads to data structures that can be analysed 8. Interoperability with major data storage or communications protocols for wearables (ideally)	1. Battery life > 48 hours 2. Battery charging time < 120 minutes 3. No location tracking function or ability to disconnect the function 4. If the device has a screen or other interface, it should not generate health-related information which can be conflicting with the ASCAPE-generated information (to prevent patient confusion) 5. Dermatologically tested 6. Comfortable to “wear” 7. Easy to care for 8. Waterproof 9. Durable

Final Required Data Items List

Table 8 Required data from ASCAPE pilot wearable

Collected data	Raw data		Aggregate data	
	Y/N	Unit of measurement	Y/N	Time level; unit of measurement
Steps	N		Y	Daily; total number of steps
Activity type	N		Y	Daily**
Activity time	N		Y	Daily; minutes
Calories burned	N		Y	Daily; kcal
Heart rate	Y*	beats per minute (bpm)	N	
Sleep quality	N		Y Sleep efficiency Wake after sleep onset	Every night; % for sleep efficiency (total sleep time/time in bed x 100%); minutes for wake after sleep onset
Sleep quantity	N		Y (total sleep time and time in bed)	Every night; hours

*Raw data to be used for investigating more specific parameters as average bpm per day, peak bpm per day, resting heart rate

**Some wearables have automated activity detection capability which seems to be reliable (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6552445/). Otherwise, we need to define the activity type based on heart rate, steps, calories burned etc. which can increase the need for raw data

Breast Cancer-Specific Requirements

Table 9 Data requirements for wearables related to data determinants for breast cancer patients' QoL monitoring

QoL Issues	Data to be collected from wearable											
	Steps	Activity type	Activity time	Calories burned	Heart rate	Heart rate variability	Oxygenation	Sleep quality	Sleep quantity	Body temperature (optional)	Gait analysis (optional)	Blood pressure (optional)
Fatigue	X	X	X	X	x	X		x	x			
Joint pain	X	X	X	X				x	x		x	
Neurotoxicity	X	X	X	X	x			x	x		x	
Anxiety	X	X	X		x	X	x	x	x			x
Depression	X	X	X	X	x			x	x			
Cognitive impairment	X	X	X	X				x	x		x	
Insomnia	X	X	X	X	x			x	x			
Lymphedema		X	X									
Hot flushes	X	X	X	X	x	X		x	x	x		
Body image	X	X	X	X				x	x			
Sexual dysfunction	X	X	X	X				x	x			
Local symptoms after surgery		X	X					x	X			
Dry vagina												
Body changes	X	X	X	X	x	X		x	x			
Emotional symptoms (loneliness)	X	X	X					x	x			

Prostate Cancer-Specific Requirements

Table 10 Data requirements for wearables related to data determinants for prostate cancer patients' QoL monitoring

QoL Issues	Data to be collected from wearable											
	Steps	Activity type	Activity time	Calories burned	Heart rate	Heart rate variability	Oxygenation	Sleep quality	Sleep quantity	Body temperature (optional)	Gait analysis (optional)	Blood pressure (optional)
Fatigue	x	X	x	X	x	X		x	x	x	x	x
Anxiety	x	X	x	X	x	X	x	x	x			x
Depression	x	X	x	X	x	X		x	x			x
Cognitive impairment	x	X	x	X				x	x		x	
Incontinence	x	X	x	X				x	x			
Lower urinary tract symptoms								x	x			x
Sexual dysfunction	x	X	x	X				x	x			
Musculoskeletal pain	x	X	x	X				x	x		x	
Weight changes	x	X	x	X	x	X		x	x		x	x
Colitis (specific for radiotherapy)	x	X	x	X	x	X		x	x			

Hot flushes	x	X	x	X	x	X		x	x	x		
Loss of libido	x	X	x	X				x	x			

Wearable Data Requirements Prioritisation

Table 11 Prioritisation of the data to be collected from wearables (breast cancer; 15 QoL Issues)

Collected data	Total N of patients' needs (n = 15)	Total N of patients' needs with priority A*	Total N of patients' needs with priority B*	Total N of patients' needs with priority C-E*
Steps	12	10	7	8
Activity type	14	16	6	1
Activity time	14	14	10	2
Calories burned	10	4	8	3
Heart rate	1	1	6	2
Heart rate variability	1	0	6	0
Oxygenation	1	0	0	2
Sleep quality	14	7	4	10
Sleep quantity	14	5	6	12
Body temperature (optional)	1	1	1	0
Gait analysis (optional)	3	2	1	1
Blood pressure (optional)	1	0	2	0

*Each collected variable was prioritised with A (highest priority) to E (lowest priority) separately by three (3) clinical partners. As a result, the total number of patients' QoL issues with priority A, B, or C-E has as maximum of 45 (1 priority from each clinical partner and total 15 patients' needs)

Table 12 Prioritisation of the data to be collected from wearables (prostate cancer; 12 QoL Issues)

Collected data	Total N of patients' needs (n = 12)	Total N of patients' needs with priority A*	Total N of patients' needs with priority B*	Total N of patients' needs with priority C-E*
Steps	11	13	11	0
Activity type	11	13	10	0
Activity time	11	14	11	0
Calories burned	11	8	13	1
Heart rate	6	2	8	4
Heart rate variability	6	0	6	6
Oxygenation	1	0	1	2
Sleep quality	12	12	10	6
Sleep quantity	12	12	10	6
Body temperature (optional)	2	0	1	0
Gait analysis (optional)	4	6	2	0
Blood pressure (optional)	5	1	4	4

*Each collected variable was prioritised with A (highest priority) to E (lowest priority) separately by three (3) clinical partners. As a result, the total number of patients' needs with priority A, B, or C-E has as maximum of 36 (1 priority from each clinical partner and total 12 patients' needs)

Evaluation of Candidate Wearables

Table 13 Evaluation of Candidate Wearables: Data Collected

Device Name	**Mi Band 4**	**Amazfit Bip S**	**Amazfit Bip**	**Vivofit 4**	**Vivosmart 4**	**Vivosport**	**Galaxy Fit**	**Charge 4**	**Inspire HR**	**Inspire**	**Band 4 Pro**	**Band 4**
Vendor	Xiaomi	Xiaomi	Xiaomi	Garmin	Garmin	Garmin	Samsung	Fitbit	Fitbit	Fitbit	Huawei	Huawei
Device Type	Activity Tracker	Smartwatch	Smartwatch	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker	Activity Tracker
Steps	X	X	X	X	X	X	X	X	X	X	X	X
Activity type	X	X	X	X	X	X	X	X	X	X	X	X
Activity time	X	X	X	X	X	X	X	X	X	X	X	X
Calories burned	X	X	X	X	X	X	X	X	X	X	X	X
Heart rate	X	X	X		X	X	X	X	X		X	X
Heart rate variability												
Oxygenation					X						X	X
Sleep quality	X	X	X	X	X	X	X	X	X		X	X
Sleep quantity	X	X	X	X	X	X	X	X	X	X	X	X
Body temperature (optional)												
Gait analysis (optional)												
Blood pressure (optional)												
Conclusion	Passed	Passed	Passed	Excluded	Passed	Passed	Passed	Passed	Passed	Excluded	Passed	Passed

Table 14 Evaluation of Candidate Wearables: Technical Committee Criteria

Device Name	Vendor	Device Type	Data Export functionality	Purchase Cost	Previous Experience[1]	Conclusion
Mi Band 4	Xiaomi	Activity Tracker	• No REST API available • Data can be exported only through the vendor's application as a Zip file (manual process) or through 3rd party applications (i.e. Google Fit)	Acceptable	Yes	Excluded due to the limitations in the data export functionality
Amazfit Bip S	Xiaomi	Smartwatch	• No REST API available • Data can be exported only through the vendor's application as a Zip file (manual process) or through 3rd party applications (i.e. Google Fit)	Acceptable	Yes	Excluded due to the limitations in the data export functionality
Amazfit Bip	Xiaomi	Smartwatch	• No REST API available • Data can be exported only through the vendor's application as a Zip file (manual process) or through 3rd party applications (i.e. Google Fit)	Acceptable	Yes	Excluded due to the limitations in the data export functionality
Vivosmart 4	Garmin	Activity Tracker	• Data can be exported via REST API upon vendor's approval • Data can be only exported in the form of daily summaries (not suitable for heart rate monitoring)	Acceptable	Yes	Excluded due to the limitation of summary formats
Vivosport	Garmin	Activity Tracker	• Data can be exported via REST API upon vendor's approval • Data can be only exported in the form of daily summaries (not suitable for heart rate monitoring)	Acceptable	Yes	Excluded due to the limitation of summary formats
Galaxy Fit	Samsung	Activity Tracker	• No REST API available • Data can be exported only through the vendor's	Acceptable	No previous experience	Excluded due to the limitations in

[1] Previous experience is considered at a vendor level, not on the specific device of the specific vendor.

			application as a Zip file (manual process)			the data export functionality
Charge 4	Fitbit	Activity Tracker	• Data can be exported via REST API upon vendor's approval	Acceptable	Yes	Passed
Inspire HR	Fitbit	Activity Tracker	• Data can be exported via REST API upon vendor's approval	Acceptable	Yes	Passed
Band 4 Pro	Huawei	Activity Tracker	• No REST API available • Data can be exported only through the vendor's application as a Zip file (manual process)	Acceptable	No previous experience	Excluded due to the limitations in the data export functionality
Band 4	Huawei	Activity Tracker	• No REST API available • Data can be exported only through the vendor's application as a Zip file (manual process)	Acceptable	No previous experience	Excluded due to the limitations in the data export functionality

Table 15 Evaluation of Candidate Wearables: Data-specific Requirements

Device Name	**Charge 4**	**Inspire HR**
Vendor	Fitbit	Fitbit
Device Type	Activity Tracker	Activity Tracker
Security, Privacy and Anonymity, on the device	Partially. An account with pseudonym is required.	Partially. An account with pseudonym is required.
Security, Privacy and Anonymity, on the cloud or repository which is retrieving data collected on the device	Partially. An account with pseudonym is required.	Partially. An account with pseudonym is required.
Collection of data related to the Patient Needs for which Predictions are made (see Table 8)	Yes (See Table 13)	Yes (See Table 13)
Data collection that is as frequent as required by the algorithms to perform their Predictions	Yes	Yes
Data collection should not stop if communications interface (i.e. Bluetooth) is unavailable	Yes	Yes
Memory for data collected and not synchronized i.e. due to communications interface being unavailable) should last at least 48 hours	Yes	Yes
Data collection that leads to data structures that can be analyzed	Yes	Yes
Interoperability with major data storage or communications protocols for wearables (ideally)	Yes	Yes

Table 16 Evaluation of Candidate Wearables: Non-Data Specific Requirements

Device Name	**Charge 4**	**Inspire HR**
Vendor	Fitbit	Fitbit
Device Type	Activity Tracker	Activity Tracker
Battery life > 48 hours	Yes	Yes
Battery charging time < 120 minutes	Yes	Yes
No location tracking function or ability to disconnect the function	Built-in GPS sensor that can be disabled.	No GPS
If the device has a screen or other interface, it should not generate health-related information which can be conflicting with the ASCAPE-generated information (to prevent patient confusion)	Partially. The device provides smart notifications, reminders or suggestions for exercise, sleep and physical activity mainly based on the goals set by the users that might be conflicting.	Partially. The device provides smart notifications, reminders or suggestions for exercise, sleep and physical activity mainly based on the goals set by the users that might be conflicting.
Dermatologically tested	Yes	Yes
Comfortable to “wear”	Yes	Yes
Easy to care for	Yes	Yes
Waterproof	Yes	Yes
Durable	Yes	Yes

Annex II. Mapping Processes

Table 17 Mapping process of QoL collection strategies and definitions within breast cancer retrospective datasets

QoL issue	Collection strategy	Questionnaire	Values	Definition for the presence of QoL issue
Anxiety BcBase Sörmland	Proxy Questionnaire	NA Fibromyalgia Impact Questionnaire	0 to 1 0 to 10	1 (defined as at least two consecutive prescriptions of anxiolytic medications) >5 (in specific questions)
Body changes	-	-	-	-
Body image Uppsala	Questionnaire	BreastQ	0 to 100	
Cognitive impairment Barcelona	Simple question	NA	0 to 1	1
Depression BcBase Sörmland	Proxy Questionnaire	NA Fibromyalgia Impact Questionnaire	0 to 1 0 to 10	1 (defined as at least two consecutive prescriptions of antidepressants) >5 (in specific questions)
Dry vagina	-	-	-	-
Emotional symptoms (loneliness) Barcelona	Simple question	NA	0 to 1	1
Fatigue Barcelona CareAcross Sörmland	Simple question Simple question Questionnaire	NA NA Fibromyalgia Impact Questionnaire	0 to 3 True/False 0 to 10	>0 True >5 (in specific questions)
Hot flushes Barcelona CareAcross	Simple question Simple question	NA NA	0 to 3 True/False	>0 True
Insomnia BcBaSe	Proxy	NA	0 to 1	1 (defined as at least two consecutive

CareAcross Sörmland	Simple question Questionnaire	NA Fibromyalgia Impact Questionnaire	True/False 0 to 10	prescriptions of medication against insomnia) True >5 (in specific questions)
Joint pain Barcelona Sörmland	Simple question Questionnaire	NA Brief Pain Inventory	0 to 3 0 to 10	>0 >5 in specific questions
Local symptoms after surgery Uppsala	Questionnaire	BreastQ	0 to 100	
Lymphedema Barcelona CareAcross Uppsala	Simple question Simple question Questionnaire	NA NA DASH	0 to 1 True/False 0 to 100	1 True
Neurotoxicity Barcelona BcBaSe CareAcross	Simple question Proxy Simple question	NA NA NA	0 to 1 0 to 1 True / False	1 1 defined as at least two consecutive prescriptions of medication against neuropathic pain True
Sexual dysfunction Barcelona Uppsala	Simple question Questionnaire	NA BreastQ	0 to 1 0 to 100	1

Table 18 Mapping process of QoL collection strategies and definitions within breast cancer prospective datasets

QoL issue	Collection strategy	Questionnaire	Values	Definition for the presence of QoL issue
Anxiety Barcelona	Questionnaire	HADS	0 to 21	(> 6 if baseline between 0-6) OR (> 10 if baseline < 10) 0-4: minimal anxiety; 5-9: mild anxiety; 10-14: moderate anxiety; 15-21: severe anxiety (> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
CareAcross	Questionnaire	GAD7	0 to 21	
Örebro	Questionnaire	HADS	0 to 21	
Body changes Örebro	Questionnaire	EORTC BR45	1 to 4 in each question	>2 in specific questions
Body image Örebro	Questionnaire	BreastQ	0 to 100	
Cognitive impairment Barcelona Örebro	Simple question Questionnaire	NA EORTC QLQ30	1 to 4 1 to 4 in each question	>1 >2 in specific questions
Depression Barcelona	Questionnaire	HADS	0 to 21	(> 6 if baseline between 0-6) OR (> 10 if baseline < 10) 0-7: minimal depression; 8-14: mild depression; 15-19: moderate-to-severe depression; 20-27: severe depression (> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
CareAcross	Questionnaire	PHQ9	0 to 27	
Örebro	Questionnaire	HADS	0 to 21	
Dry vagina Barcelona	Simple question	NA	1 to 4	>1

Örebro	Questionnaire	EORTC BR45	1 to 4 in each question	>2 in specific questions
Emotional symptoms (loneliness) Barcelona	Questionnaire	Three-Item Loneliness Scale	1 to 3	
Fatigue Barcelona CareAcross Örebro	Simple question Simple question Questionnaire	NA NA EORTC QLQ30	1 to 3 True/False 1 to 4 in each question	>1 True >2 in specific questions
Hot flushes Barcelona CareAcross Örebro	Simple question Simple question Questionnaire	NA NA EORTC BR45	1 to 4 True/False 1 to 4 in each question	>1 True >2 in specific questions
Insomnia Barcelona CareAcross Örebro	Simple question Simple question Questionnaire	NA NA EORTC QLQ30	1 to 4 True/False 1 to 4 in each question	>1 True >2 in specific questions
Joint pain Barcelona Örebro	Simple question Questionnaire	NA EORTC BR45	1 to 4 1 to 4 in each question	>1 >2 in specific questions
Local symptoms after surgery Örebro	Questionnaire	BreastQ	0 to 100	
Lymphedema Barcelona CareAcross Örebro	Simple question Simple question Questionnaire	NA NA DASH	1-40 True/False 0 to 100	1 True >40
Neurotoxicity Barcelona CareAcross Örebro	Simple question Simple question Questionnaire	NA NA EORTC BR45	0 to 1 True/False 1 to 4 in each question	1 True >2 in specific questions
Sexual dysfunction Örebro	Questionnaire	EORTC BR45	1 to 4 in each question	>2 in specific questions

Table 19 Mapping process of QoL collection strategies and definitions within prostate cancer retrospective datasets

QoL issue	Collection strategy	Questionnaire	Values	Definition for the presence of QoL issue
Anxiety CareAcross	Questionnaire	GAD7	0 to 21	0-4: minimal anxiety; 5-9: mild anxiety; 10-14: moderate anxiety; 15-21: severe anxiety
Bowel dysfunction CareAcross Örebro	Simple question Questionnaire	NA NA	True/False 1 to 4	True >2 in specific questions
Cognitive impairment	-	-	-	-
Depression CareAcross	Questionnaire	PHQ9	0 to 27	0-7: minimal depression; 8-14: mild depression; 15-19: moderate-to-severe depression; 20-27: severe depression
Erectile dysfunction CareAcross Örebro	Simple question Questionnaire	NA IIEF-5	True/False 0 to 25	True 1 - 7 severe; 8-11 moderate
Fatigue CareAcross	Simple question	NA	True/False	True
Hot flushes CareAcross	Simple question	NA	True/False	True
Incontinence CareAcross	Simple question	NA	True/False	True
Low urinary tract symptoms CareAcross Örebro	Simple question Questionnaire	NA IPSS	True/False 0 to 25	True 8-18 moderate; > 18 severe
Loss of libido	-	-	-	-
Musculoskeletal pain	-	-	-	-
Weight changes CareAcross	Simple question	-		drop of >5% in six months

Table 20 Mapping process of QoL collection strategies and definitions within prostate cancer prospective datasets

QoL issue	Collection strategy	Questionnaire	Values	Definition for the presence of QoL issue
Anxiety				
Athens	Questionnaire	HADS	0 to 21	(> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
CareAcross	Questionnaire	GAD7	0 to 21	0-4: minimal anxiety; 5-9: mild anxiety; 10-14: moderate anxiety; 15-21: severe anxiety
Örebro	Questionnaire	HADS	0 to 21	(> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
Bowel dysfunction				
Athens	Questionnaire	EORTC PR25	1 to 4 in each question	>2 in specific questions
CareAcross	Simple question	NA	True/False	True
Örebro	Questionnaire	EORTC PR25	1 to 4 in each question	>2 in specific questions
Cognitive impairment				
Athens	Questionnaire	EORTC QLQ30	1 to 4 in each question	>2 in specific question
Örebro	Questionnaire	EORTC QLQ30	1 to 4 in each question	>2 in specific questions
Depression				
Athens	Questionnaire	HADS	0 to 21	(> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
CareAcross	Questionnaire	PHQ9	0 to 27	0-7: minimal depression; 8-14: mild depression; 15-19: moderate-to-severe depression; 20-
Örebro	Questionnaire	HADS	0 to 21	

				27: severe depression (> 6 if baseline between 0-6) OR (> 10 if baseline < 10)
Erectile dysfunction Athens CareAcross Örebro	Questionnaire Simple question Questionnaire	IIEF NA IIEF	0 to 30 Domain A; 0 to 10 Domain B; 0 to 10 Domain C; 0 to 15 Domain D; 0 to 10 Domain E True/False 0 to 30 Domain A; 0 to 10 Domain B; 0 to 10 Domain C; 0 to 15 Domain D; 0 to 10 Domain E	Domain A < 14 OR Domain B < 5 OR Domain C < 5 OR Domain D < 8 OR Domain E < 5 True Domain A < 14 OR Domain B < 5 OR Domain C < 5 OR Domain D < 8 OR Domain E < 5
Fatigue Athens Örebro	Questionnaire Questionnaire	EORTC QLQ30 EORTC QLQ30	1 to 4 in each question 1 to 4 in each question	>2 in specific question >2 in specific questions
Hot flushes Athens CareAcross Örebro	Questionnaire Simple question Questionnaire	EORTC PR25 NA EORTC PR25	1 to 4 in each question True/False 1 to 4 in each question	>2 in specific questions True >2 in specific questions
Incontinence Athens CareAcross Örebro	Questionnaire Simple question Questionnaire	EORTC PR25 NA EORTC PR25	1 to 4 in each question True/False 1 to 4 in each question	>2 in specific questions True >2 in specific questions
Low urinary tract symptoms Athens CareAcross Örebro	Questionnaire Simple question Questionnaire	EORTC PR25 NA EORTC PR25	1 to 4 in each question True/False 1 to 4 in each question	>2 in specific questions True >2 in specific questions
Loss of libido Athens Örebro	Questionnaire Questionnaire	IIEF IIEF	0 to 10 for Domain C 0 to 10 for	Domain C < 5 Domain C < 5

			Domain C	
Musculoskeletal pain				
Athens	Questionnaire	EORTC QLQ30	1 to 4 in each question	>2 in specific question
Örebro	Questionnaire	EORTC QLQ30	1 to 4 in each question	>2 in specific questions
Weight changes				
Athens	Measurements	NA		Drop of > 10% from baseline
CareAcross	Simple question	NA		Drop of >5% in six months
Örebro	Measurements	NA		Drop of > 10% from baseline

References

[1] AEAT, "Agencia Tributaria," 2017. [Online]. Available: https://tinyurl.com/y2xkunlg. [Accessed July 2020].

[2] M. Perdomo, C. Davies, K. Levenhagen and K. Ryans, "BREAST CANCER EDGE TASK FORCE OUTCOMES: Assessment Measures of Secondary Lymphedema in Breast Cancer Survivors," *Rehabilitation Oncology,* vol. 32, no. 1, pp. 22-35, 2014.

[3] S. Harrington, S. Miale and D. Ebaugh, "Breast Cancer EDGE Task Force Outcomes: Clinical Measures of Health Related Quality of Life," *Rehabilitation Oncology,* vol. 33, no. 1, pp. 5-17, 2015.

[4] S. E. Harrington, L. Gilchrist, J. Lee, F. Westlake and A. Baker, "Oncology Section EDGE Task Force on Cancer: A Systematic Review of Clinical Measures for Pain," *Rehabilitation Oncology,* vol. 36, no. 2, pp. 83-92, 2016.

[5] M. Alappattu, S. E. Harrington, A. Hill, A. Roscow and A. Jeffrey, "Oncology Section EDGE Task Force on Cancer: A Systematic Review of Patient-Reported Measures for Sexual Dysfunction," *Rehabilitation Oncology,* vol. 35, no. 3, pp. 137-143, 2017.

[6] S. Harrington, J. Lee, G. Colon and M. Alappattu, "Oncology Section EDGE Task Force on Prostate Cancer A Systematic Review of Outcome Measures for Health-Related Quality of Life," *Rehabilitation Oncology,* vol. 34, no. 1, pp. 27-35, 2016.

[7] "Wearable Technology in Healthcare - A Thematic Research," Global Data UK Ltd, 2019.

[8] A. Mahajan, G. Pottie and W. Kaiser, "Transformation in Healthcare by Wearable Devices for Diagnostics and Guidance of Treatment," *ACM Transactions on Computing for Healthcare,* vol. 1, no. 1, pp. 1-12, 2020.

[9] International Panel on Climate Change (IPCC), Climate Change 2001: Synthesis Report. A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Integovernmental Panel on Climate Change, R. T. Watson and Core Writing Team, Eds., Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press, 2001, p. 398.

[10] R. C. Estoque, T. Togawa, M. Ooba, K. Gomi, S. Nakamura, Y. Hijioka and Y. Kameyama, "A review of quality of life (QOL) assessments and indicators: Towards a "QOL-Climate" assessment framework," *Ambio,* vol. 48, no. 6, pp. 619-638, 2019.

[11] R. Jayadevappa, R. Cook and S. Chhatre, "Minimal important difference to infer changes in health-related quality of life-a systematic review," *Journal of Clinical Epidemiology,* vol. 89, pp. 188-198, 2017.

[12] M. T. King, "A point of minimal important difference (MID): a critique of terminology and methods," *Expert Review of Pharmacoeconomics & Outcomes*

Research, vol. 11, no. 2, pp. 171-184, 2011.

[13] F. Cardoso, S. Kyriakides, S. Ohno, F. Penault-Llorca, P. Poortmans, I. T. Rubio, S. Zackrisson and E. Senkus, "Early Breast Cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Annals of Oncology,* vol. 30, no. 8, pp. 1194-1220, 2019.

[14] C. Parker, E. Castro, K. Fizazi, A. Heidenreich, P. Ost, G. Procopio, B. Tombal and S. Gillessen, "Prostate cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up," *Annals of Oncology,* vol. 31, no. 9, p. 2020, 2020.

[15] N. Mottet, J. Bellmunt, M. Bolla, E. Briers, M. G. Cumberbatch, M. D. Santis, N. Fossati, T. Gross, A. M. Henry, S. Joniau, T. B. Lam, M. D. Mason, V. B. Matveev, P. C. Moldovan and v. d. B, "EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent," *European Eurology,* vol. 71, no. 4, pp. 618-629, 2017.

[16] "Oxford Centre for Evidence-based Medicine - Levels of Evidence," [Online]. Available: https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/.

[17] J. Han, M. Kamber and J. Pei, Data Mining: Concepts and Techniques, Elsevier, 2011.

[18] I. H. Witten, E. Frank and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 2011.

[19] G. James, D. Witten, T. Hastie and R. Tibshirani, An Introduction to Statistical Learning, Springer, 2013.

[20] A. Donabedian, "Evaluating the quality of medical care.," *Milbank Q.,* vol. 83, no. 4, pp. 691-729, 1966.

[21] A. L. Cochrane, Effectiveness and efficiency: Random reflections on health services., Nuffield Trust, 1972.

[22] P. Aspden, J. M. Corrigan, J. Wolcott and S. M. Erickson, Eds., Patient safety: achieving a new standard for care, Washington, DC, 2004.

[23] M. Gray, "Designing healthcare for a different future," *Journal of the Royal Society of Medicine,* vol. 109, no. 12, pp. 453-458, 2016.

[24] T. Weldring and S. M. S. Smith, "Article Commentary: Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs)," *Health Services Insights,* 2013.

[25] B. W. S. R. G. S. Manary MP, "The patient experience and health outcomes," *N Engl J Med.,* vol. 368, no. 3, pp. 201-203, 2013.

[26] N. Black, "Patient reported outcome measures could help transform healthcare," *BMJ,* vol. 346, pp. 1-5, 2013.

[27] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly,* vol. 13, no. 3, p. 319–340, 1989.

[28] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes,* vol. 50, no. 2, p. 179–211, 1991.

[29] B. J. Fogg, "A behavior model for persuasive design," *InProceedings of the 4th*

international Conference on Persuasive Technology, pp. 1-7, 2009.

[30] "MAFEIP Site," [Online]. Available: https://www.mafeip.eu/. [Accessed 15 August 2020].